From DEPARTMENT OF NEUROBIOLOGY,
CARE SCIENCES AND SOCIETY
Karolinska Institutet, Stockholm, Sweden

SICKNESS ABSENCE AMONG PATIENTS WITH CHRONIC PAIN IN SWEDISH SPECIALIST HEALTHCARE

Riccardo LoMartire

Stockholm 2021

Published by Karolinska Institutet.
Printed by Universitetsservice US-AB, 2021

ISBN 978-91-8016-435-1

Sickness Absence Among Patients With Chronic Pain In Swedish Specialist Healthcare

THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Riccardo LoMartire

The thesis will be defended in public at the Erna Möller hall, Karolinska Institutet Neo, Blickagången 16, Huddinge on Friday the 14th of January 2022 at 10.00 AM.

Principal Supervisor:
Professor Björn O. Äng
Region Dalarna
Department of Research
and Higher Education

Co-supervisors:
Dr Linda Vixner
Dalarna University
School of Health and Welfare

Professor Björn Gerdle
Linköping University
Department of Health,
Medicine and Caring Sciences
Pain and Rehabilitation Centre

Opponent:
Professor Anders Kottorp
Malmö University
Faculty of Health and Society
Department of Care Science

Examination Board:
Professor Kristina Alexanderson
Karolinska Institutet
Department of Clinical Neuroscience
Division of Insurance Medicine

Associate professor Erin Gabriel
Karolinska Institutet
Department of Medical Epidemiology
and Biostatistics

Professor Anne Söderlund
Mälardalen University
School of Health, Care and Social Welfare
Division of Physiotherapy

To Aurora, Nathaniel, and Élodie.

"A good traveler has no fixed plans
and is not intent on arriving."

Lao Tzu

ABSTRACT

Background: Chronic pain beyond three months is a global public health problem. Every third adult suffers from a chronic pain condition, resulting in a socioeconomic burden that corresponds to 3-10% of gross domestic product in western economies. This burden can be largely attributed to absenteeism-related productivity loss where a few highly impaired individuals are the most resource-intensive. Simultaneously, a detailed overview of sickness absence (SA) associated with chronic pain is complicated by incongruent classification due to conflicting perspectives on the condition as either a symptom or a disease in its own right.

Aim: Based on a well-defined chronic pain population in the Swedish specialist healthcare, this thesis primarily aims to provide a SA overview, to explore the possibility of SA prevention, and to evaluate interdisciplinary treatment (IDT) as a SA intervention. A secondary objective was to assess the psychometric properties of three questionnaires that measure the core domains of the chronic pain experience.

Methods: The aims were addressed in three register-based studies using microdata from five Swedish national registers. Study I used sequence analysis to describe SA in 44,241 patients over a 7-year period and subsequently developed a machine learning-based model to predict chronic pain-related SA in the final two years. Study II emulated a target trial to compare the total SA duration over a 5-year period for 25,613 patients that were either included in an IDT program or in other/no interventions. Study III analyzed the properties of the Short Form-36 Health Survey (SF-36), the EuroQol 5-Dimensions instrument (EQ-5D), and the Hospital Anxiety and Depression Scale (HADS) within the item response theory-framework.

Results: SA increased from 17% to 48% over the five years before specialist healthcare entry to then decrease to 38% over the final two years. With information on eight predictors, it was possible to discriminate between patients that would have low or high SA in the coming two years with 80% accuracy. SA trends were similar for patients in IDT programs and other/no interventions, albeit the IDT patients had 67 (95% CI: 48, 87) more SA days over the complete 5-year period. Finally, the psychometric evaluation revealed that SF-36 adequately captured physical and mental health, while HADS was suitable as a measure of overall emotional distress, and EQ-5D had insufficient precision for any meaningful application.

Conclusion: Our findings are most useful to guide policy and research. SA in the studied patients remained high over the entire observation period. Decision support tools could prove valuable in identifying patients at risk of high SA earlier in the healthcare chain in order to direct preventative measures. We found no support for IDT decreasing SA more than other/no interventions, but it is possible that this was a consequence of our methodology. Further studies of the IDT effects are needed, but uncontrolled designs that attribute SA change over time to IDT are inappropriate for this purpose, as the SA peak observed around specialist healthcare entry is likely to be driven by the referral procedure. Finally, SF-36 and HADS are psychometrically sound measures of the chronic pain experience core domains.

PREFACE

This doctoral process was like wandering in an unfamiliar landscape. It involved unexpected struggles but was also greatly rewarding, and the destination from the onset was unknown. What began as an inquiry into the field of pain became increasingly centered on research methodology. Misalignment between expert opinion and evidence triggered the question: How certain can we be of our knowledge?

In chronic pain rehabilitation, experts take the superiority of interdisciplinary treatment (IDT) over less comprehensive interventions for granted, and conceptually, it is difficult to understand why this would not be the case. It is intuitive that IDT would be more effective: based on holistic biopsychosocial theory, it organizes transdisciplinary resources in a coordinated attack on the different facets of chronic pain. How could a treatment that adapts a multiplicity of other treatments to the problem at hand not be the best solution? Theory does not always correspond to practice, however, and despite an ever-increasing body of research, the scientific literature is equivocal for many of the outcomes for IDT. Perhaps this is a consequence of crude methodology combined with the inherent complexity of the field of study. On the other hand, it is no secret that healthcare resources are limited and that important theoretical components may be lost in economically feasible practice. The question as to whether IDT is an effective treatment strategy thereby remains.

This thesis could not provide a definite answer to this question, but instead raised a larger question: How should we organize our resources to avoid evidence stagnation?

LIST OF SCIENTIFIC PAPERS

I. LoMartire R, Dahlström Ö, Björk M, Vixner L, Frumento P, Constan L, Gerdle B, Äng BO. Predictors of sickness absence in a clinical population with chronic pain. J Pain. 2021;22(10):1180-1194.

II. LoMartire R, Björk M, Dahlström Ö, Constan L, Frumento P, Vixner L, Gerdle B, Äng BO**.** The value of interdisciplinary treatment for sickness absence in chronic pain: a nationwide register-based cohort study. Eur J Pain. 2021. 25(10):2190-2201.

III. LoMartire R, Äng BO, Gerdle B, Vixner L. Psychometric properties of Short Form-36 Health Survey, EuroQol 5-Dimensions, and Hospital Anxiety and Depression Scale in patients with chronic pain. Pain. 2020;161(1):83-95.

CONTENTS

ABBREVIATIONS

ATC	Anatomical Therapeutic Chemical Classification System
CI	Confidence interval
EQ-5D	EuroQol 5-Dimensions
HADS	Hospital Anxiety and Depression Scale
HRQoL	Health-related Quality of Life
IASP	International Association for the Study of Pain
ICD	International Classification of Diseases
IDT	Interdisciplinary treatment
IRT	Item response theory
LISA	Longitudinal Integration Database for Health Insurance and Labour Market Studies
MiDAS	Micro Data for Analysis of the Social Insurance Register
NPR	National Patient Register
PDR	Prescribed Drug Register
RMSEA	Root mean square error of approximation
SA	Sickness absence including both sick leave and disbility pension
SEK	Swedish crowns
SF-36	Short Form-36 Health Survey
SQRP	Swedish Quality Registry for Pain Rehabilitation
SRMSR	Standardized root mean square residual
WHO	World Health Organization

1 INTRODUCTION

1.1 CHRONIC PAIN

Classification

Pain needs no presentation, yet it is elusive. The International Association for the Study of Pain (IASP) defines pain as:[147]

> "*An unpleasant sensory and emotional experience associated with, or resembling that associated with, actual or potential tissue damage.*"

Pain is thereby a personal experience that extends beyond a physical sensation to encompass other domains of life. When pain persists over a longer period of time, it can transform from an essential survival mechanism into the harmful state of chronic pain.[32]

Chronic pain is a controversial condition. More than half a century has passed since its maladaptive destructiveness was first acknowledged.[13] Yet, the perception that it is merely a symptom of another disease persists until today.[65] A paradigm shift is taking place towards recognizing chronic pain as a disease in its own right, and it was recently added to the 11th revision of the World Health Organization's (WHO) international classification system of diseases (ICD-11) under the diagnostic code MG30.[128,146,183,197] Consistent with the common interpretation of IASP's earlier definition, chronic pain is defined by WHO as:[121,197]

> *"Pain that persists or recurs for longer than 3 months."*

This new proposal distinguishes between chronic primary pain referring to conditions where it is considered a disease in itself and chronic secondary pain where it originates from other diseases.[127,183] The chronic primary pain definition highlights the central roles of emotional distress and functional disability in the condition:

> *"Pain in one or more anatomical regions that (1) persists or recurs for longer than 3 months, (2) is associated with significant emotional distress (eg, anxiety, anger, frustration, or depressed mood) and/or significant functional disability (interference in activities or daily life and participation in social roles), (3) and the symptoms are not better accounted for by another diagnosis."*

It can be subclassified into: chronic widespread pain, complex regional pain syndrome, chronic primary headache or orofacial pain, chronic primary visceral pain, and chronic primary musculoskeletal pain.[127,183] Chronic secondary pain is classified into chronic cancer-related pain, chronic postsurgical or posttraumatic pain, chronic neuropathic pain, chronic secondary headache or orofacial pain, chronic secondary visceral pain, or chronic secondary musculoskeletal pain.[183] This thesis mainly involves chronic primary pain and chronic secondary musculoskeletal and headache pain.

Societal burden

Chronic pain is a major public health problem. Approximately a third of the global general population is believed to suffer from a chronic pain condition.[32,49,91,106,165] Prevalence varies widely with pain definition and is unevenly distributed so that it increases with female sex, age, and low socioeconomic status.[32,53,106,165] The Global Burden of Disease study has monitored time lived with disability due to non-fatal conditions since 1990, and currently covers 369 diseases and injuries in 204 countries and territories globally.[60] Musculoskeletal disorders are consistently identified among the leading causes and constituted 17% of all the time lived with non-fatal disability in 2019.[63] Low back pain alone accounted for 7% of the total burden and has invariably been isolated as the single leading condition since 1990.[63] Other important contributors linked to chronic pain are found among mental and neurological disorders, and include migraine, major depression, and anxiety disorders each representing 5%, 4%, and 3% of the total burden in 2019, respectively (Figure 1).[63] It is virtually impossible to estimate the extent of this burden that can be attributed to chronic pain specifically, but the condition likely represents a considerable proportion. While most people recover quickly from a pain episode, the prognosis rapidly deteriorates with persistent pain.[8] As many as two-thirds of low back pain patients in the primary healthcare reportedly transit from acute to chronic pain, with a majority still having pain after one year.[85,90] Cost evaluations of different pain conditions also consistently support that a small proportion of individuals with the highest disability represent the majority of the total burden.[74,75,80,106,114]

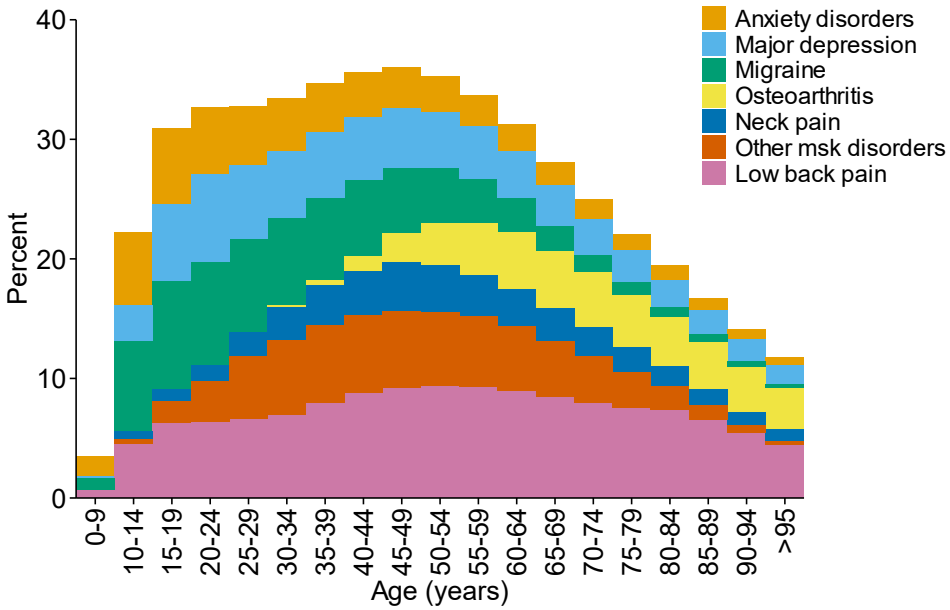

Figure 1. Percent of total time lived with disability per age group for selected non-fatal condition linked to chronic pain in 2019. Source: Global Burden of Disease Collaborative Network.

Socioeconomically, chronic pain represents 3–10% of the annual gross domestic product in western economies.[19,57,139] In Sweden, the annual costs were estimated at 87.5 billion SEK in 2003, while annual costs for diagnoses linked to chronic pain were reported at €32 billion for 2008.[70,171] Most investigations support that indirect costs due to productivity loss from absenteeism account for the bulk of the total costs.[19,70,139,171,187] Conditions associated with chronic pain are the principal causes of prolonged sick leave and disability pension, with musculoskeletal disorders, mental disorders, and injuries consistently included among the leading diagnoses.[4,78,109,188] Musculoskeletal and mental disorders associated with chronic pain are also the leading causes of sickness absence in Sweden, accounting for 68% of public health insurance costs in 2009.[173,174] However, despite numerous studies indirectly linking chronic pain to an enormous socioeconomic burden in the form of absenteeism, research targeting chronic pain specifically is surprisingly scarce.

Etiology

The human can be conceptualized as an adaptive supersystem made up of nested and interdependent subsystems that strive towards sustaining homeostasis (the balance of essential survival processes).[25,66] All adaptive systems share the critical features of irritability, connectivity, and plasticity, meaning that they are dynamic and responsive to perturbation, their components interact, and they can phase shift in response to environmental alternations.[25] When pain disrupts homeostasis, it induces an allostatic stress response (a coordinated regulatory defense mechanism) in neural, endocrine, and immune subsystems.[25] Succinctly, this response can be divided into an immediate component that operates via the sympathetic-adreno-medullar axis to activate short-term fight-flight-freezing behaviors and a slower component that operates via the hypothalamus-pituitary-adrenal axis to mobilize resources for long-term defence.[25,64] The allostatic response is a resource-intensive process that under normal circumstances is followed by recovery.[25] Prolonged stress and dysregulation (failure to recover) in any of the subsystems or their connectivity ultimately disturbs the entire supersystem and introduces a perpetual phase shift.[25] With chronic pain, the pathophysiology is not fully understood, but the ability to respond and recover from perturbations is disturbed and pain modulating systems are dysfunctional.[25] Central and, to a lesser extent, peripheral sensitization are considered important mechanisms that lead to hypersensitivity, hyperalgesia, and allodynia of pain due to increased neural excitability and reduced endogenous pain inhibition.[105]

The neuromatrix theory of pain postulates that pain is produced by the neuromatrix (a widely distributed neural network) in the brain.[119,120] The neuromatrix consists of neurons between the thalamus, the cortex, and the limbic system, which involves sensory-discriminative, affective, and cognitive perception.[120] The chronic pain experience is determined by neurosignatures (characteristic nerve impulse patterns) that vary with genetic predisposition and sensory input.[120] In turn, sensory input is influenced by physical stimuli, psychological

stress, and social environment.[120] The pain experience and its associated behaviors is thus dependent on a dynamic interplay between biopsychosocial factors.[50]

Experience

Chronic pain can be a debilitating condition that permeates all aspects of life. What begins as a pain sensation, may with time result in impairments that cause more suffering than the sensation itself. Whereas individual trajectories are highly variable, their central components can be conceptualized within the physical, emotional, and social domains (Figure 2). Physical disability often interferes with everyday activities such as walking, exercising, household chores, and work.[18,45,141] Sleep disturbances are also common and may express themselves as poor quality, reduced duration, or insomnia.[7,99] Emotional distress in the form of anxiety, depression, or anger is another fundamental co-morbidity.[21,41,88,148] Both emotional distress and sleep disturbances have a bidirectional relationship to the pain, meaning that they exacerbate each other in a negative feedback loop.[6,32,88] Over time, this physical and emotional deterioration adversely effects relationships and even increases the risk of employment loss and divorce, which may progressively lead to social deprivation due to a gradual withdrawal from society.[18,32,45,135,141] Negative consequences further extend to surrounding family and friends, whom may suffer greatly due to the additional everyday burden, strained interactions, and financial stress.[45,141] Typical expressions also include physical struggle, mental exhaustion, hopelessness, social isolation, and a general perception of life being impoverished and confined.[182] All these factors combined contribute to decreased health-related quality of life (HRQoL), which reportedly can be as poor in individuals with chronic pain as in those with terminal cancer.[6]

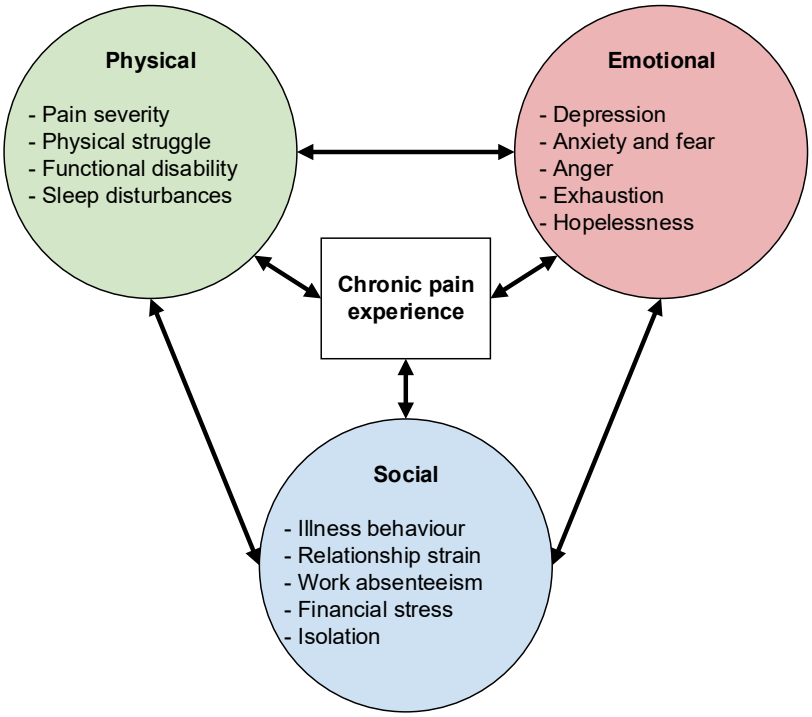

Figure 2. Schematic overview of the chronic pain experience.

The chronic pain experience is notoriously difficult to measure and considerable resources have been invested into isolating its core domains.[29,48,95,183,185] This is an ongoing process, with domains still overlapping, morphing, and amalgamating; however, three central domains that consistently recur across different entities over time are pain severity, emotional distress, and physical function.[29,48,95,183,185] Unlike characteristics such as weight and height, human experiences are unobservable latent traits that are inferred from indicators through statistical procedures.[151] Self-report questionnaires are the primary source of information for evaluating the chronic pain experience in the clinical setting. For them to accurately capture a latent trait, it is necessary that they are valid in the population and setting where they are applied, which implicates that they are based on solid underlying theory and have statistically robust empirical properties.[151] Three widely used questionnaires in the health sciences are the Short Form-36 Health Survey (SF-36) of physical and mental health, EuroQol 5-Dimensions (EQ-5D) of HRQoL, and the Hospital Anxiety and Depression Scale (HADS) of emotional distress.[20,77,193,198] Although they have been recommended specifically for chronic pain, research has simultaneously highlighted the need for evaluation of their psychometric properties as support for their validity is insufficient in this population.[26,30,33,48]

Intervention

IASP classifies pain interventions into four categories according to their complexity:[178]

(1) *unimodal treatments:* single interventions that target specific pain mechanisms.
(2) *multimodal treatments:* several concurrently administered interventions from within a single discipline that target different pain mechanisms.
(3) *multidisciplinary treatment:* several concurrently administered interventions from within multiple disciplines that target different pain mechanisms.
(4) *interdisciplinary treatment:* several concurrently administered interventions involving a multidisciplinary team that collaborates in the assessment and treatment of patients based on the biopsychosocial model.

It is generally recognized that the biopsychosocial perspective is the most appropriate in the assessment and treatment of chronic pain.[32,58,59] At the highest level of complexity, interdisciplinary treatment (IDT) is considered a chronic pain core intervention.[58,96] It is theoretically grounded in the biopsychosocial model of illness, which postulates that the natural course depends on a dynamic interaction between biopsychosocial factors.[32,50,59] By concurrently directing physical, psychological, and social measures at the different facets of pain, intervention components are believed to act both independently and in conjunction with each other.[58,96] Because the chronic pain experience is person-specific, it is recommended that IDT programs are adapted to each patient's individual need. Common modalities include physical therapy, acceptance and commitment therapy, cognitive behavioral therapy, occupational therapy, and pharmaceutical treatment.[58,96] However, the practical implementation of IDT is highly variable due to personalization and local

differences in expertise. IDT fulfills the criteria of complex interventions, which are notoriously difficult to evaluate due to their comprehensive and flexible composition intended to target multifaceted behaviors.[36] This provides an explanation for the limited evidence in support of IDT, despite its strong theoretical foundation and the numerous studies conducted to evaluate its effects.[44,160,171,172,186] Whereas research suggests that IDT is marginally superior to less comprehensive interventions where the chronic pain experience is concerned, evidence for its effects on sickness absence are inconclusive.[44,160,171,172,186]

In Swedish specialist healthcare, IDT is offered to patients, the majority of whom are referred from primary healthcare, by approximately 40 clinics nationwide.[153] National guidelines imply that intervention is administered in cohesive programs by experienced interprofessional teams that collaborate to personalize care.[61] In practice, IDT programs vary across clinics, but they primarily focus on group-based activities such as cognitive behavioral therapy, physical exercise and occupational training, which are concurrently administered by a team of physicians, physiotherapists, occupational therapists, and social workers.[61,153] Programs are typically delivered 2–5 days per week over a 4-12-week period and consist of more than 30 hours in total.[153] Not everyone who visits a specialist clinic is included in an IDT program, however. As a rule, patients are initially evaluated by the team and may be assigned to other/no interventions depending on the evaluation's outcome, their own preferences, and other unspecified factors. Indicators that influence assignment to an IDT program are sickness absence history, pain interference with everyday activities, emotional distress, and confidence in recovery, while age, sex, socioeconomic status, and policy are known to affect healthcare in general.[2,15,39,61,116,123]

1.2 SICKNESS ABSENCE

Sickness absence (SA), or absence from work due to incapacity, is a multifactorial phenomenon that is determined at the structural levels of a society, organization, and individual.[3,179] Diverse theories in the fields of psychology, sociology, economics, and medical science have presented overlapping yet different perspectives that, either directly or indirectly, pertain to the underlying causes of SA.[5] Individual characteristics that influence SA include age, sex, socioeconomic status, and disability, organizational factors include satisfaction due to economic, social, and psychological factors, and societal factors refer to macroeconomics, labor market conditions, and social norms.[3,164]

The Swedish social insurance system has been implemented to provide economic stability in case of work incapacity. The typical working age in Sweden is 18 to 65 years with retirement possible from the age of 61. All Swedish residents aged 16 to 64 with minimum income from employment or unemployment are eligible for social insurance benefits if their ability to work is reduced due to disease or illness. Sick leave is possible from age 16 and is granted as full or part time of ordinary work hours.[132] Spells are generally compensated for by the employer the first 14 days, including a qualifying period with no compensation the

first day.[132] From the 15th day, spells are reimbursed and recorded by the Swedish Social Insurance Agency.[132] Since 2008, sick leave benefits are restricted to 364 days per 450-day period, with exceptions made for serious illness or if a person's ability to work has been reduced due to an occupational injury.[132] Disability pension is also granted at either full or part time, with a permanent pension option available for individuals aged 30-64 when their working capacity is deemed permanently impaired, and a temporary form possible for individuals aged 19-29.[132]

1.3 CONCEPTS OF SCIENCE

Causality

Causality is the relationship between cause and effect. In essence, it can be viewed from two opposite perspectives of nature as either inherently deterministic or stochastic.[136] The former represents the deterministic Laplacian view where randomness reflects an imperfect understanding of nature's underlying mechanisms, while the latter corresponds to the stochastic view of modern quantum mechanics where determinism reflects simplification to facilitate human understanding.[136] This thesis adapts a quasi-deterministic view of causality, contextualized through counterfactuals that are consolidated with graph theory, structural causal models, and stochastic mechanisms.[81,101,107,136] In this context, causal inference is an intrinsically epistemic concept that requires prior knowledge to interpret data causally. The graph theory provides a visual framework for the assumptions of the overall causal structure (Figure 3A), the structural models express the relationships mathematically via sets of structural equations that represent causal mechanisms, and the stochastic component reflects human's uncertainty about nature.[69,136] However, it is the counterfactuals that ultimately allow us to differentiate between causation and association. Counterfactuals are hypothetical scenarios that contrast potentialities with the purpose to derive information on ontological causal relationships.[136] They can be conceptualized under the multiverse and are epistemic through their limitation to modal knowledge (i.e., metaphysical necessities and possibilities in terms of 'must', 'could', and 'could not') under our perceived laws of nature.[107,196] Hence, in the context of causal inference, counterfactuals are restricted to the possible worlds.

The ideal randomized experiment represents the best available emulation of the counterfactual multiverse. Properly conducted random assignment ascertains that the average sample characteristics are probabilistically equivalent across experimental groups, which implies that observed and unobserved biases are balanced.[161] Under optimal conditions, the experimental groups are thus exchangeable and can be conceptualized as mirror images of parallel worlds in all but the intervention.[81] In ideal randomized experiments, the intention-to-treat effect of the intervention assignment is identical to the as-treated effect of the intervention itself (Figure 3B).[81] However, in practice, they differ with the influence of systematic attrition, non-adherence, and non-concealment.[81] Imperfect randomized experiments are subject to the same type of biases as observational studies and require

situation-specific evaluation to determine whether the intention-to-treat or as-treated principle is the most appropriate.[84,115] The causal structure of as-treated effects in randomized experiments corresponds to that of the observational design and is thereby susceptible to confounding (Figure 3B). Likewise, both the intention-to-treat and the as-treated principles in experiments with differential attrition that are restricted to available participants can induce bias via non-causal pathways (Figure 3C). In light of the discrepancy between theory and practice of the randomized experiment, it has been argued that the observational design can be a better choice in pragmatic evaluation of clinical interventions.[71]

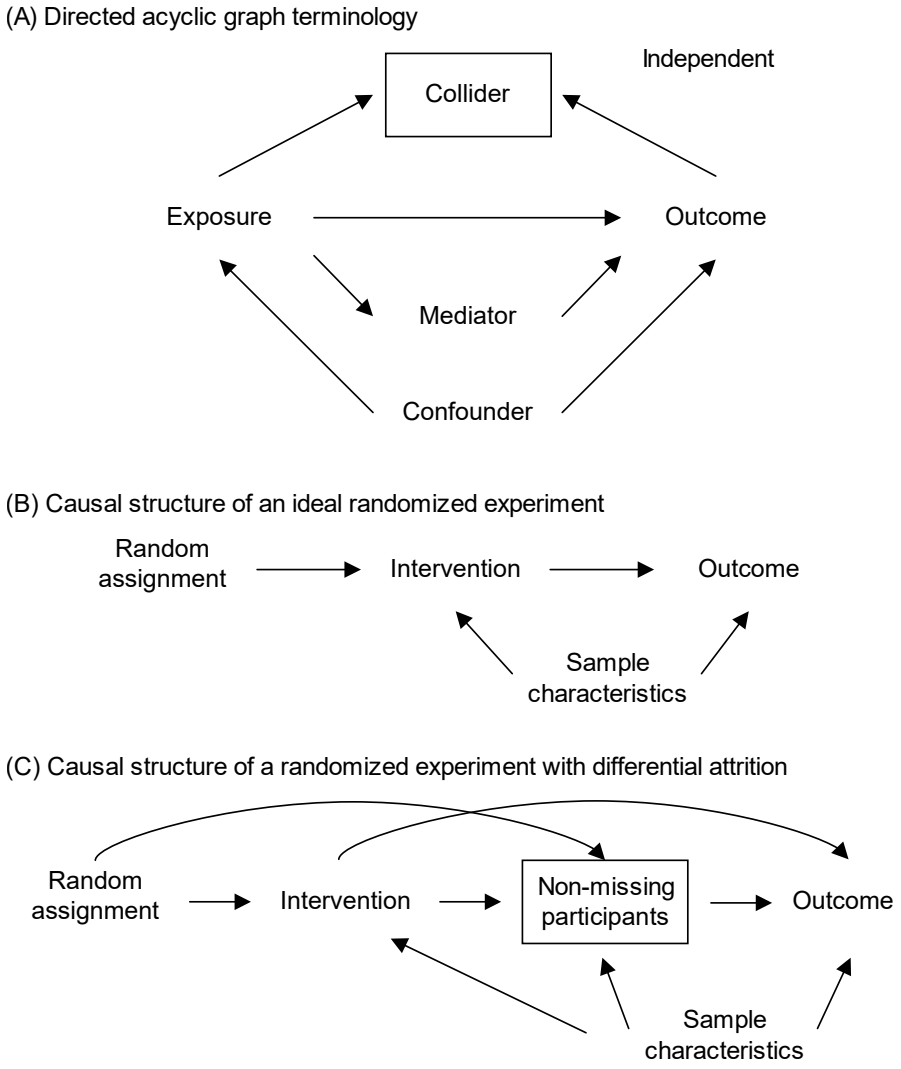

Figure 3. Directed acyclic graphs (DAG) of causal structures. (A) DAG terminology and rules. The pathways *Exposure* → *Outcome* and *Exposure* → *Mediator* → *Outcome* represent direct and indirect causal effects, respectively. The pathways *Exposure* ← *Confounder* → *Outcome* and *Exposure* → *Collider* ← *Outcome* represent non-causal pathways via a common cause and a common effect, respectively. Unconditioned pathways via mediators and confounders are open, while pathways via colliders are closed until conditioned on (denoted by rectangle). *Independent* represents a variable with no relationships to other variables. (B) Ideal randomized experiment with concealed assignment, no attrition, and perfect compliance. The intention-to-treat effect of the assignment *Random assignment* → *Intervention* → *Outcome* is identical to the as-treated effect of the intervention itself *Intervention* → *Outcome* and there are no non-causal pathways from assignment to outcome. (C) Unblinded randomized experiment with differential attrition. Both the intention-to-treat and as-treated effects are susceptible to bias as the conditioning on non-missing participants opens a non-causal pathway via sample characteristics to the outcome.

Target trial emulation is a recommended strategy for observational studies of causal intervention effects.[81,83] In this strategy, the protocol of the observational study is adapted so that it corresponds to a hypothetical randomized experiment (Table 1).[81,83] Target trial emulation requires a sufficiently large dataset of adequate quality for meaningful causal inferences and is therefore often most suitable for national register studies. Practically, individuals that meet the study eligibility criteria are divided into groups based on the received intervention and subsequently contrasted with respect to an outcome of interest. To ensure that differences in the outcome reflect the causal effect of the intervention, all non-causal pathways must be closed. This requires strong assumptions on causal relationships based on prior knowledge and is managed through any of several methods (i.e., stratification, regression adjustment, matching, inverse probability weighting, standardization, or a combination of these). Finally, it is critical that the starting point is well-defined so that the conditions for all intervention groups under evaluation are the same.

Table 1. Target trial emulation protocol adapted from Hernan and Robins 2016.

Component	Description	Example
Eligibility criteria	Defines the target population that the results are generalizable to via inclusion and exclusion criteria.	Adults with non-cancer pain for more than 90 days that visited a specialist clinic between 2009 and 2016.
Intervention strategies	Describes the intervention and control.	Either an interdisciplinary treatment program at a specialist healthcare clinic or physical therapy in the primary healthcare.
Assignment procedure	Defines the causal structure and the methods used to close non-causal pathways.	Inverse probability weights are used to adjust for baseline confounders identified in the scientific literature
Observation period	Defines the starting point (time zero), the maximum duration of the follow-up, and other end points (e.g., death or loss to follow-up).	Time zero is the first visit to the clinic and patients are followed up to maximum five years.
Outcome	Defines the outcome under evaluation (e.g., death or sick leave).	Total days of sickness absence over the five-year period from time zero.
Causal contrast	Specifies the causal question of interest (e.g., intention-to-treat, as-treated, or per-protocol).	The intention-to-treat effect of intervention assignment.
Analysis plan	Specifies the practical analyses used to derive the numeric results.	Poisson regression with inverse probability weights for each patient was used to regress sickness absence on intervention.

Prediction

Prediction is the statement of an expected outcome given available information. It is innately non-causal, as its focus lies on the correct forecast and the underlying data generating mechanisms are of secondary importance. Conceptually, the objective of predictive models is to learn the functional relationship that links the features (predictor candidates) to the outcome from the data. The learning process requires an algorithm to describe the relationship, evaluation criteria to define success, and an optimization method to search for the relationship that maximizes success.[76] Prediction problems can be divided into either regression of quantities or classification of labels.[76] As a general rule, predictive performance and interpretability increases and decreases with algorithm complexity, respectively; however, there is no universally superior algorithm to all problems and performance is instead problem-specific.[17,76] A useful predictive model is sufficiently complex to capture the patterns of interest, while being applicable to datasets not used in the model development. This concept is known as the bias-variance trade-off and refers to the balance of the systematic and random components in the prediction error, which is needed to optimize generalizability (Figure 4A).[76] In practice, predictive performance is typically optimized in a cross-validation procedure where data is split into parts that, independently from each other, are used to learn patterns, evaluate performance, and finally confirm the performance evaluation (Figure 4B).[76]

Latent variables

Latent variables are unobservable constructs that must be inferred from observed indicators.[14] They can represent human experiences such as depression that are measured by self-report questionnaires and societal phenomena such as socioeconomic status that are created from register data. The former is defined as a reflective latent variable, which influences its indicators, and the latter as a formative latent variable, which is determined by its indicators (Figure 5).[46,47,54] Their difference carries conceptual and practical implications. A reflective latent variable exists independently of its indicators, which, in turn, reflect the latent variable's status, covary because of it, and are unrelated without it. In contrast, a formative latent variable does not necessarily represent a real construct, but instead summarizes the information of its indicators, which may not covary as they correspond to different causes (dimensions) of the construct. The conceptual framework of latent variables is thus critical in their interpretation.

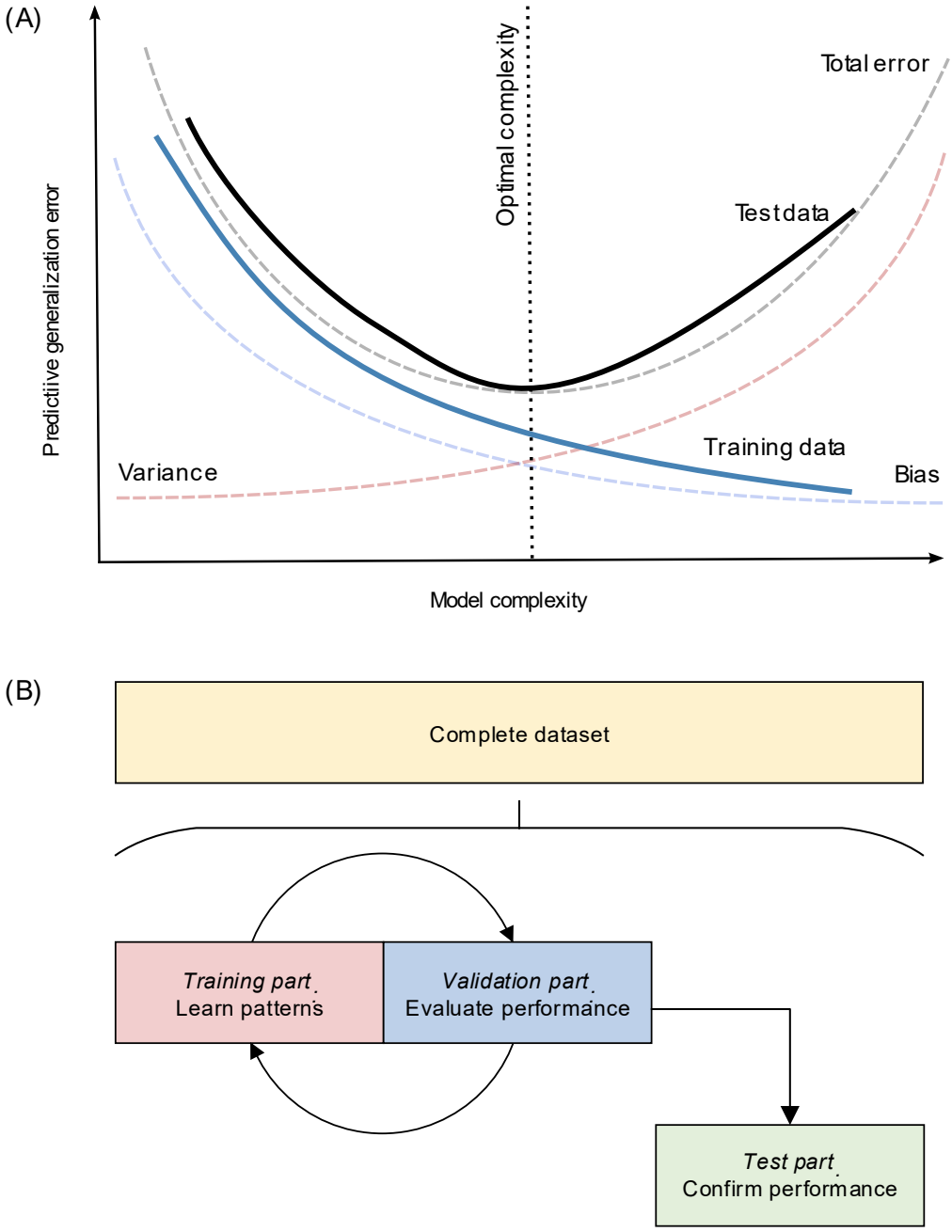

Figure 4. (A) Conceptual overview of the bias-variance trade-off inspired by Hastie et al., 2017. The predictive generalization error consists of three components: two reducible bias and variance components that increase and decrease with model complexity, respectively, as more parameters enable patterns of higher complexity but are simultaneously estimated with less precision; and an irreducible variance component related to the sampling uncertainty. A too simplistic model fails to capture the general patterns of interest (underfit), while a too complex model captures sample-specific patterns and thus generalizes poorly too new data (overfit). (B) Overview of the cross-validation procedure. With data randomly split into three parts, predictive performance is optimized in an iterative procedure where the algorithm is trained to learn patterns on the first part and performance is evaluated on the second part. Once the final model is selected, its performance is confirmed on the third part to avoid indirect overfit.

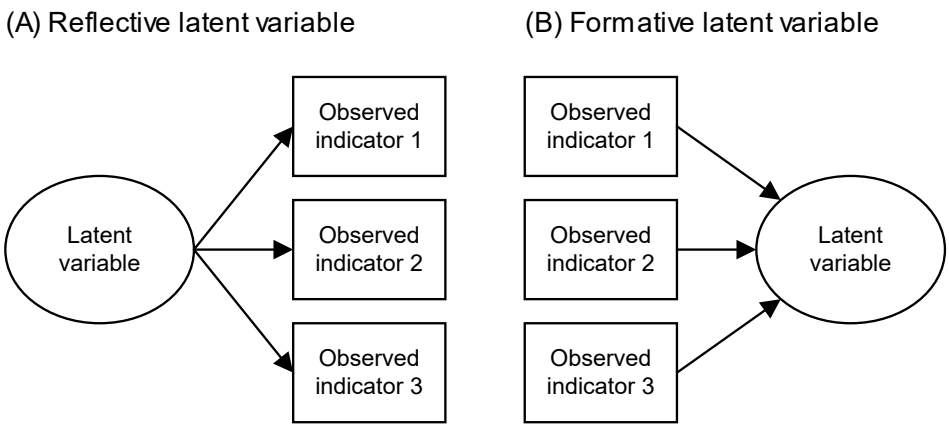

Figure 5. Schematic overview of a reflective and a formative latent variable inspired by Borsboom 2003.

1.4 RATIONALE

Chronic pain is a globally prevalent condition that represents an enormous socioeconomic burden. This burden can largely be attributed to SA-related productivity loss where a few highly impaired individuals account for a disproportionate amount of the costs. However, a detailed SA overview is complicated by incongruent chronic pain classification due to conflicting perspectives on the condition as either a symptom or a disease. Most SA studies target associated musculoskeletal disorders rather than chronic pain itself, which is problematic, as chronic pain entails other dimensions than the physical, while musculoskeletal disorders are not restricted to chronic-pain diagnoses. It is thereby uncertain to what extent the results of such studies are specific to chronic pain, which increasingly is gaining recognition as its own disease.

In addition, optimal pain management remains a conundrum, despite the numerous chronic pain interventions existing today, with treatment effects known to be both inconsistent and small. IDT is a theoretically appealing and internationally recommended core intervention, but the current state of evidence is inconclusive where its effects on SA are concerned. This can largely be attributed to the combined complexity of chronic pain and IDT, which is recognized as a major impediment for high-quality studies. In the absence of properly designed and conducted randomized controlled trials, large-scale observational studies are the next best alternative for effect evaluations under counterfactual causal inference.

To meaningfully interpret the SA of the population under study, it is necessary to characterize their chronic pain experience. The subjective and heterogeneous nature of the chronic pain experience makes it notoriously difficult to capture, but emotional distress and physical function are recognized as core domains. These constructs are primarily measured through self-report questionnaires, with SF-36, EQ-5D, and HADS recommended specifically for individuals with chronic pain. However, their psychometric properties are insufficiently evaluated in this population and are in need of further investigation to avoid biased measurement.

2 RESEARCH AIMS

The primary objectives of this thesis were to provide a sickness absence overview, to explore the possibility of sickness absence prevention, and to evaluate the effects of interdisciplinary treatment on sickness absence among chronic pain patients in Swedish specialist healthcare. Because these questions are contingent on the chronic pain experience, indicators of which have been insufficiently evaluated, we also aimed to assess whether its core domains were adequately captured by available questionnaires. These objectives were addressed in three studies:

- **Study I.** To describe sickness absence and explore the possibility of predicting future high sickness absence at entry into specialist healthcare.

- **Study II.** To evaluate the effect of interdisciplinary treatment as a sickness absence intervention.

- **Study III.** To assess the psychometric properties of the Short Form-36 Health Survey, the instrument EuroQol 5-Dimensions, and the Hospital Anxiety and Depression Scale.

3 METHODS

3.1 DESIGN AND PARTICIPANTS

This thesis consists of three register-based nationwide observational studies of chronic pain patients in the Swedish specialist healthcare system (Table 3). The source population was patients in the Swedish Quality Register for Pain Rehabilitation (SQRP), characterized by complex clinical presentations and non-response to primary care interventions, and corresponding to roughly 0.5‰ of the Swedish population annually. Table 2 details the sample characteristics.

Study I was a cohort study that described SA from five years before to two years after entry into specialist care and developed a model for classifying patients based on their SA in the final two years. It included 44,241 patients who made their first visit to a clinic during the period 2005 to 2016 (82.4% of the source population). Study II emulated a target non-blinded randomized controlled trial to estimate the population-average effects of IDT on SA over a five-year period. It included 25,613 patients (47.1% of the source population) who had visited a specialist clinic between 2009 and 2016. Study III was a cross-sectional psychometric evaluation of three self-report questionnaires that capture core domains of the chronic pain experience. It included 35,908 adult cancer-free SQRP patients (84.6% of the source population) who had visited a specialist clinic between 2009 and 2016. Studies I and II included data from the Micro Data for Analysis of the Social Insurance Register, the Longitudinal Integration Database for Health Insurance and Labour Market Studies, the National Patient Register, and the Swedish Prescribed Drug Register, whereas study III was based strictly on SQRP data.

Table 2. Study sample characteristics.

	Study I	Study II		Study III
		Interdisciplinary treatment	*Other/no interventions*	
Patients[a]	44,241 (100.0)	13,628 (100.0)	11,985 (100.0)	35,908 (100.0)
Demographics				
Age (years)[b]	44 (36, 52)	42 (34, 49)	42 (33, 50)	45 (36, 53)
Female[a]	31,610 (71.4)	10,247 (75.2)	8,018 (66.9)	25,744 (71.7)
Born in Sweden[a]	34,374 (77.7)	10,861 (79.7)	8,981 (74.9)	27,843 (77.5)
University/college education (>12 years)[a]	11,881 (26.9)	4,274 (31.4)	3,249 (27.1)	10,048 (28.0)
Employed[a]	28,525 (64.5)	10,237 (75.1)	7,529 (62.8)	22,853 (63.6)
Family's past 5-year mean annual disposable income (1000 SEK)[b]	174 (129, 233)	194 (148, 250)	179 (131, 235)	189 (141, 247)
Disability				
Pain duration (years)[b]	5.7 (2.1, 12.5)	4.6 (1.7, 11.2)	4.8 (1.9, 11.0)	5.4 (1.9, 12.7)
NRS-10 past week pain intensity[b]	7 (6, 8)	7 (6, 8)	7 (6, 8)	7 (6, 8)
Most prevalent ICD-10 diagnoses[a]				
Fibromyalgia (M79.7)	6,026 (13.6)	2,320 (17.0)	1,668 (13.9)	5,532 (15.4)
Unspecified pain (R52.9)	3,515 (7.9)	1,121 (8.2)	1,353 (11.3)	3,382 (9.4)
Myalgia (M79.1)	3,299 (7.5)	1,067 (7.8)	1,101 (9.2)	2,843 (7.9)
Low-back pain (M54.5)	3,150 (7.1)	1,249 (9.2)	880 (7.3)	2,793 (7.8)
Cervicobrachial syndrome (M53.1)	2,426 (5.5)	861 (6.3)	630 (5.3)	1,986 (5.5)
HADS emotional distress[b]	47 (33, 60)	48 (35, 60)	47 (33, 61)	47 (33, 60)
EQ-5D index[b]	0.2 (0.0, 0.6)	0.2 (0.0, 0.6)	0.2 (0.0, 0.6)	0.2 (0.0, 0.6)
MPI interference[b]	4.6 (3.8, 5.3)	4.5 (3.7, 5.2)	4.5 (3.6, 5.3)	4.5 (3.8, 5.3)
High confidence in recovery[a]	7,691 (17.4)	3,209 (23.5)	2,063 (17.2)	6,410 (17.9)
Sickness absence[c]				
Gross past-year sick leave days[b]	59 (0, 273)	89 (0, 260)	39 (0, 258)	38 (0, 238)
Ongoing sick leave at first visit to clinic[a]	19,536 (44.2)	6,984 (51.3)	4,906 (40.9)	14,963 (41.7)
Ongoing permanent disability pension at first visit to clinic[a]	7,791 (17.6)	0 (0)	0 (0)	5,195 (14.5)

[a] Frequency (percent). [b] Median (25^{th}, 75^{th} percentile). [c] due to chronic pain-related ICD-10 diagnoses: M(00−99), G(43−44, 47, 50−64, 82, 96−97), R(07, 10, 26, 29, 51−52), S(12−13, 22, 32, 42−43, 53), T(85, 88, 91−94), and F(32−33, 41, 43, 45).

Table 3. Study overview.

Study	Aim	Design	Registers	Observation period	Eligibility criteria	Sample size	Statistical analysis
I	Describe and predict sickness absence	Register-based cohort study	SQRP, MiDAS, LISA, NPR, PDR	7 years	*Inclusion:* ▪ IDT startyear: 2005-2016, ▪ Age at the IDT assessment: 18-64, ▪ Pain duration at the IDT assessment: ≥ 90 days *Exclusion:* ▪ A cancer diagnosis in the previous 5 years	44,241	Unsupervised models: ▪ Sequence analysis, ▪ Monothetic divisive hierarchical clustering Supervised classification models: ▪ Multinomial logistic regression, ▪ Support vector machine, ▪ Gradient boosting machine, ▪ Artificial neural network
II	Evaluate the effects of IDT on sickness absence	Register-based cohort study	SQRP, MiDAS, LISA, NPR, PDR	5 years	*Inclusion:* ▪ IDT start year: 2009-2016, ▪ Age at the IDT assessment: 18-60, ▪ Pain duration at the IDT assessment: ≥ 90 days *Exclusion:* ▪ A cancer diagnosis in the previous 5 years ▪ An IDT assessment the previous 2 years ▪ Any permanent disability pension in the previous year	25,613	Reversible Markov multistate model: ▪ Non-parametric transition model, ▪ Flexible parametric transition model with inverse probability weights ▪ Logistic regression exposure model for the inverse probability weights
III	Evaluate the psychometric properties of SF-36, EQ-5D, HADS	Register-based cross-sectional study	SQRP	Not relevant	*Inclusion:* ▪ IDT start year: 2009-2016, ▪ Age at the IDT assessment: ≥ 18, ▪ Pain duration at the IDT assessment: ≥ 90 days *Exclusion:* ▪ A cancer diagnosis at the IDT assessment	35,908	Multidimensional item response theory: ▪ Logistic graded response model ▪ Two-parameter logistic model

LISA, Longitudinal Integration Database for Health Insurance and Labour Market Studies. MiDAS, Micro Data for Analysis of the Social Insurance Register. NPR, National Patient Register. PDR, Swedish Prescribed Drug Register. SQRP, Swedish Quality Register for Pain Rehabilitation.

3.2 DATA SOURCES

Swedish National Registers

The Swedish tradition of register-keeping can be traced back to the 17th century.[55] Today, there are more than 100 national registers that are available for research purposes. They constitute an immense source of microdata that can be linked via personal identification numbers held by all Swedish citizens.[112] Registers can be categorized into central government registers and healthcare quality registers. Government registers are part of the Swedish infrastructure and routinely collect data from everyone who meets their eligibility criteria, with mandatory participation.[55] Their coverage is, therefore, typically very high. In contrast, quality registers contain information related to specific health-related conditions and operate on a voluntary basis.[55] With optional participation for individual health care units and informed consent required from individual patients, coverage is not necessarily known and could be considerably lower. In this thesis, data were included from five national registers and linked per patient via their personal identification number (Figure 6; Table 4).[112]

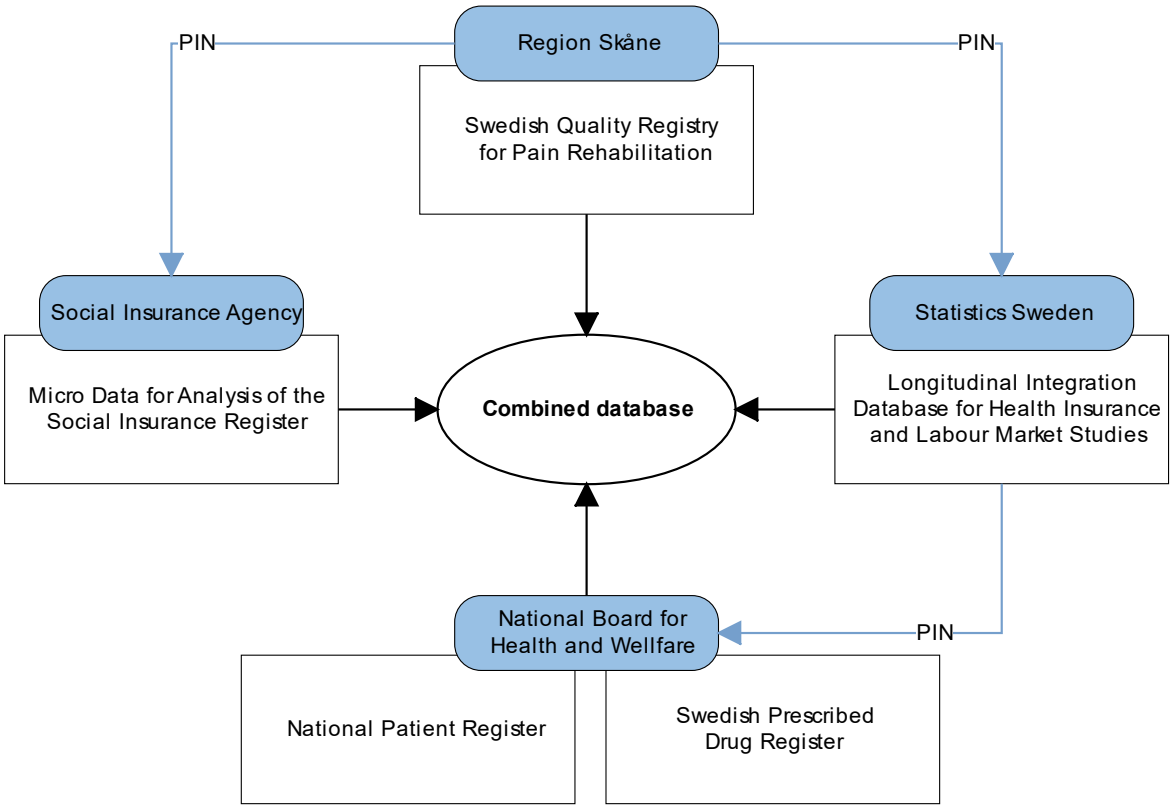

Figure 6. Data source overview. Black arrows mark data flow while blue arrows mark the flow of personal identification numbers (PIN) for data linkage.

Table 4. Data source summary.

Data source	Data source target population	Main thesis variables	Temporal resolution	Data acquisition month	Data range
SQRP	Patients with non-cancer chronic pain referred to a Swedish IDT specialist clinic	Pain duration, pain intensity, number of pain locations, main pain location, ICD-10 diagnosis, confidence in recovery, SF-36, EQ-5D, HADS, MPI, clinic's geographical region, date of first visit to clinic, IDT selection status	Up to three time points: (1) first visit to clinic (2) IDT completion, (3) 1 year after IDT completion	June 2018	January 2005 to December 2016
MiDAS	All individuals with one sick leave spell > 14 days or disability pension.	Spell dates for sick leave and disability pension, sickness absence extent, primary ICD-10 diagnosis, employment status	Daily	February 2019	January 2000 to October 2018
LISA	All individuals ≥ 16 years	Sex, age, country of birth, family composition, education level, disposable income	Annual	March 2019	December 2000 to December 2016
NPR	All individuals with ≥ 1 visit to the inpatient healthcare or outpatient specialist healthcare	Inpatient care admission and discharge date, outpatient specialist care visit date, ICD-10 diagnosis	Daily	February 2020	January 2000 to December 2018
PDR	All individuals with ≥ 1 dispensed prescriptive drug	Dispense date, ATC-code, defined daily dose	Daily	February 2020	July 2005 to December 2019

LISA, Longitudinal Integration Database for Health Insurance and Labour Market Studies. MiDAS, Micro Data for Analysis of the Social Insurance Register. NPR, National Patient Register. PDR, Swedish Prescribed Drug Register. SQRP, Swedish Quality Register for Pain Rehabilitation.

Swedish Quality Registry for Pain Rehabilitation

The Swedish Quality Registry for Pain Rehabilitation (SQRP) has consecutively aggregated data from specialist interdisciplinary pain management clinics since 1998.[131,150] The register is managed by Region Skåne via Uppsala Clinical Research Center and is temporally divided into several databases, of which SQRP-Access (1998-2009) and SQRP-1 (2007-2017) were used here. SQRP targets patients with non-cancer chronic pain who have been referred to any of the approximately 40 specialist clinics across Sweden (Figure 7) by primary health care, private general practitioners, or hospitals. Specialist healthcare is offered by both university hospitals and private healthcare providers. The register contains demographic data, patient-reported questionnaire data on the pain experience, and data reported by the care providers on ICD-10 diagnoses. During the period evaluated in this thesis, SQRP covered roughly 80-95% of the specialist clinics across Sweden; however, the participation rate of individual patients was not recorded.

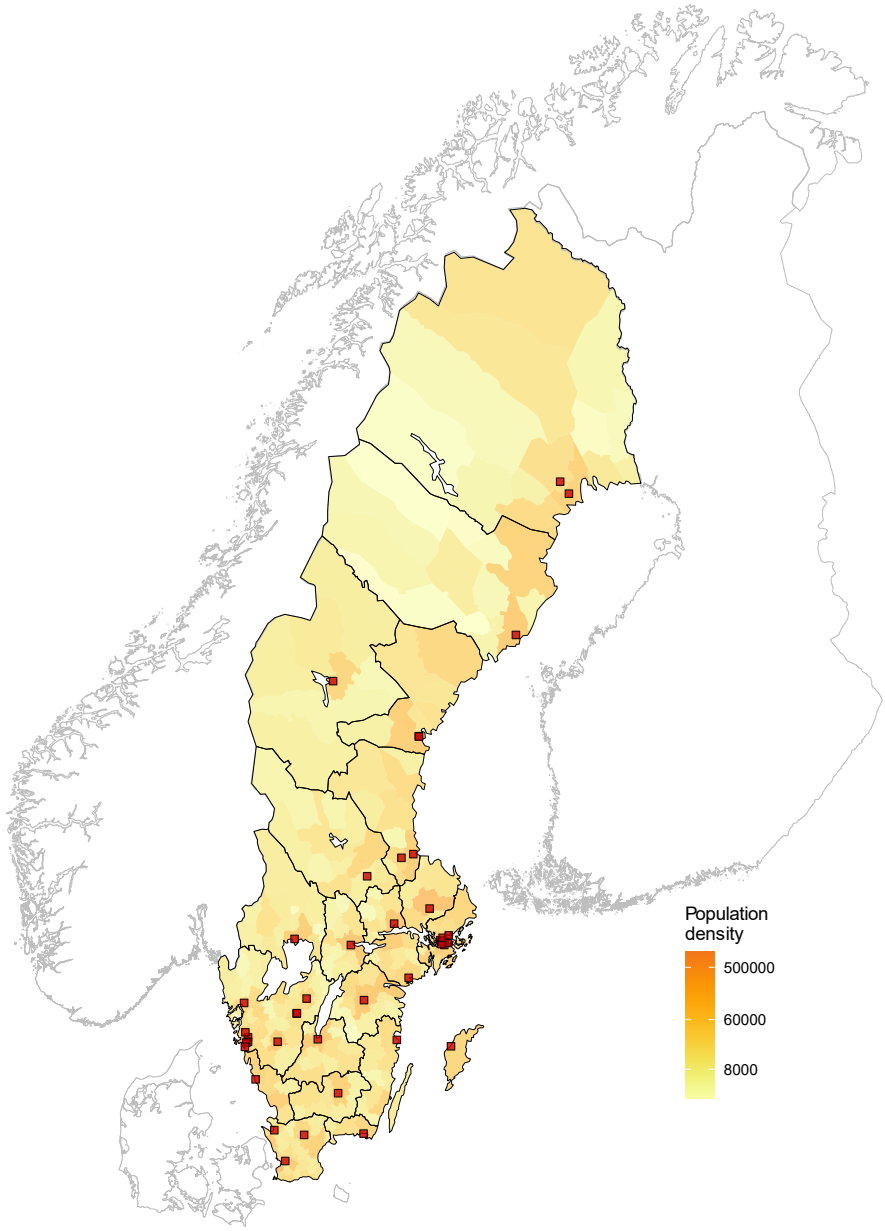

Figure 7. Distribution of Swedish SQRP clinics overlaid onto a map with population density per municipality for 2016 (Statistics Sweden). Map produced with permission from the Database of Global Administrative Areas (GADM).

Micro Data for Analysis of the Social Insurance Register

The Micro Data for Analysis of the Social Insurance Register (MiDAS) was developed in 2004-2005 to permit analyses of individual SA trajectories.[132] It was based on the database STORE, which covers decisions and reimbursements on the Swedish population from 1992.[132] MiDAS is administered by the Swedish Social Insurance Agency and contains information on sick leave spells > 14 days and disability pension spells from all individuals. More specifically, it covers spell dates, extent, ICD-10 diagnostic codes, employment status, and monetary reimbursement.

Longitudinal Integration Database for Health Insurance and Labour Market Studies

The Longitudinal Integration Database for Health Insurance and Labour Market Studies (LISA) is a meta-register established in 1990 that targets all Swedish inhabitants aged ≥ 16 years.[113,163] The register is administered by Statistics Sweden and combines data on an annual basis from primary National registers to facilitate analyses of relationships between employment and health. LISA contains information on sex, age, country of birth and citizenship, highest attained education level, and disposable income.

National Patient Register

The National Patient Register (NPR) collects data on inpatient healthcare since 1964 and outpatient specialist healthcare since 2001.[111,125] The inpatient data has complete national coverage since 1987, with 99% of somatic and psychiatric inpatient discharges registered.[111] The outpatient data coverage ranges from 70% at the inception to 97% in the recent years.[125] NPR is managed by the National Board of Health and Welfare and contains data on healthcare domain, ICD-10 diagnose codes, and date of healthcare. The validity of ICD-10 diagnoses is reportedly high, with positive predictive values of 85-95% for most diagnoses.[111]

Swedish Prescribed Drug Register

The Swedish Prescribed Drug Register (PDR) records data on dispensed prescription drugs from Swedish pharmacies since 2005.[124,190,194] Prescription drugs are estimated at 84% of total drugs dispensed and 77% of total expenditure, with prescription-free drugs and inpatient healthcare-administered drugs not covered by the register.[194] PDR is managed by the National Board of Health and Welfare and contains pharmaceutical descriptors, including product names, ATC-codes, volumes, costs, and dates dispensed.

3.3 SICKNESS ABSENCE

In this thesis, SA encompassed both sick leave and disability pension. Chronic pain-related SA contained the ICD-10 diagnoses for musculoskeletal system diseases (M: 00−99), nervous system diseases (G: 43−44, 47, 50−64, 82, 96−97), unclassified pain symptoms (R: 07, 10, 26, 29, 51−52), injuries and complications (S: 12−13, 22, 32, 42−43, 53; T: 85, 88, 91−94), and mental disorders (F: 32−33, 41, 43, 45).[70] Finally, to distinguish between temporary and permanent SA, we further combined the temporary form of disability pension for individuals below 30 years with sick leave.[3]

3.4 STATISTICS

All analyses were conducted in the open-source software R and Python (libraries are declared in supplementary Tables 1S-2S).[43,144,145,189] Table 3 details the principal statistical methods used.

In study I, we described SA and developed a prediction model of future chronic pain-related SA. Sequence analysis was used to describe SA over a seven-year period and to sort the patients according to their chronic pain-related SA patterns during the final two years.[56,169] Briefly, sequence analysis is a non-parametric technique that compares temporal series of categorical states between patients and generates a matrix of patient sequence differences.[152] To divide patients into SA subgroups, the difference matrix was analyzed with divisive monothetic hierarchical clustering.[56] Practically, it is a tree-based method that sequentially divides the patients into smaller groups; allowing cluster solutions to be compared with respect to statistical quality, robustness, and pertinence.[27,152] Predictors of the SA clusters were then identified in a machine learning-based modeling procedure with four parallel run classification algorithms.[28,62,138] The algorithms included multinomial logistic regression, a support vector machine, gradient tree boosting, and a multilayer perceptron artificial neural network.[76] The former is a parametric technique and the latter three can be conceptualized as non-parametric techniques given that their parameter number is determined by the data. In total, we explored 101 candidate predictors from the domains of sociodemographics, the chronic pain experience, SA history, and healthcare. Identified predictors were compared between algorithms and the algorithm with the highest predictive performance was selected as the final model. Balanced accuracy was used as a primary measure of overall performance, while sensitivity and the positive predictive values were included to assess the performance of individual classes.[10] To maximize generalizability, predictors were chosen in a nested cross-validation procedure, during which models were trained and evaluated on independent datasets.[76]

In study II, a target non-blinded randomized controlled trial was emulated to estimate the population-average intention-to-treat effect of IDT on SA. That is, we contrasted all patients being offered IDT with all patients being offered other/no interventions (controls). Under a set of theory-based causal assumptions, a hypothetical scenario was simulated where patients

assigned to an IDT program and to control interventions were equivalent (Figure 8).[81] Markov multistate models were used to estimate the daily probability of SA and the total SA duration over a five-year period from the first visit to a SQRP specialist clinic.[31,40] Succinctly, such models generalize the standard survival approach to account for both recurrent and competing events; however, they require that possible states, possible transitions, and transition models are selected in advance.[37,118,195] Markov models describe stochastic processes under the assumption that the next state transition is dependent on the transition history through the current state only.[118] We included three states: no SA, full or partial sick leave, and full or partial disability pension. Bidirectional transitions were possible between the former two states and unidirectional transitions were possible from any of the former two states to the disability pension state. Transitions were modeled with the non-parametric Aalen-Johansen estimator for the unadjusted observed SA and with flexible parametric survival models for population-average SA.[1,37,67] Concisely, the Aalen-Johansen estimator is a generalization of the non-parametric Kaplan-Meier estimator for the hazard of a survival process to multiple states, whereas the flexible parametric model of proportional hazards captures the shape of the baseline hazard function with restricted cubic splines that are joined at knots predefined in number and placement.[1,37,67] Restricted cubic splines are piecewise third-degree polynomials that are constrained to form smooth functions, which are linear beyond the boundary knots.[176] We used the Bayesian information criterion to guide the number of knots and placed them equally over the quantiles of the event time distribution.[176] The flexible parametric model is advantageous in that it provides a measure of absolute effects and easily incorporates time-dependent effects via interaction between covariates and splines of time.[176] To account for our causal assumptions, the parametric transition models were estimated with stabilized inverse probability weights derived from a logistic regression exposure model.[81] Conceptually, the inverse probability weights are used to construct a pseudo-population by re-weighting the patients according to their propensity score of intervention assignment, which eliminates the association between the confounders and the intervention.[81] At their theoretical boundaries, the expected weights are equal for all patients with random assignment, while they are infinitely large for the control patients with completely deterministic assignment. Finally, stabilized weights use the marginal probability of intervention assignment in the numerator to avoid extreme weights due to a low probability of intervention assignment, which can cause numerical problems.[81]

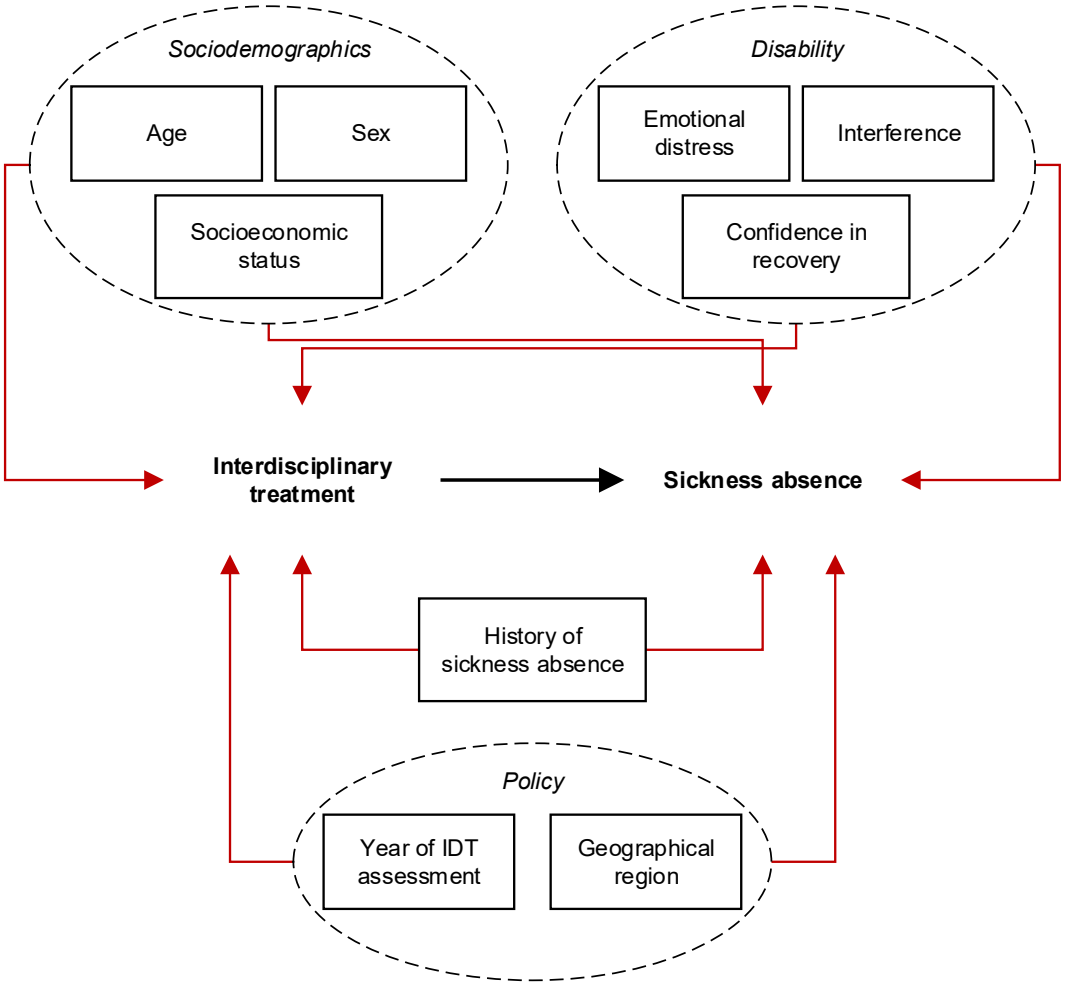

Figure 8. Assumed causal structure. Figure reproduced from study II under CC BY 4.0.

In study III, we evaluated the psychometric properties of SF-36, EQ-5D, and HADS. Analyses first targeted structural validity, then internal consistency reliability, and finally convergent/discriminant validity. Structural validity of recognized theoretical constructs for each questionnaire was examined within the item response theory (IRT)-framework.[24] Concisely, IRT models describe the relationship between observed indicators (e.g., the questionnaire items) and a stipulated latent trait (e.g., depression).[149,151] The two types of IRT models used here were the logistic graded response model for ordered multicategory item responses and the logistic two-parameter model for dichotomous item responses.[159] Unidimensional IRT models were used when theories motivated a single latent trait and corresponding multidimensional extensions otherwise. To identify the empirical representations with the best structural properties, several fit indices were compared in a five-fold cross-validation procedure.[76] These included the root mean square error of approximation (RMSEA) and the standardized root mean square residual (SRMSR) as primary global fit indicators, with estimates ≤ 0.05 considered acceptable for both.[117,151] Once the most structurally sound models were identified, the internal consistency reliability of the latent traits was computed as a summary measure of their precision, with estimates ≥ 0.70 considered acceptable.[180] Finally, convergent and discriminant validity were evaluated via the intercorrelation of latent trait scores.

3.5 ETHICS

This project was approved by Uppsala's Regional Ethical Review Board (DNR 2018/036) and conducted in accordance with both the declaration of Helsinki and the guidelines on good research practice by the Swedish Research Council. The autonomy of the patients was respected as participation in SQRP mandated a written informed consent, which declared their legal rights and the possibility of data linkage to other registers. The risk of privacy infringements was minimized by encoding personal identification numbers before data extraction from register holders, by storing patient information in a secure location, and by limiting presented results to aggregate level to prevent indirect identification. Finally, the project added no additional burden to the healthcare system as it was strictly based on routinely collected data.

4 RESULTS

4.1 SICKNESS ABSENCE OVERVIEW (STUDY I)

Between 17% and 49% of the patients received chronic pain-related SA benefits at any given time during the 7-year observation period (Figure 9). Whereas chronic pain-related SA peaked around entry into specialist healthcare, SA linked to other diagnoses remained relatively constant over the full period. Chronic pain-related SA was unevenly distributed between patients so that a fifth received no benefits, another fifth accounted for 56% of the net days, and three-fifths of the patients shared the remaining SA days.

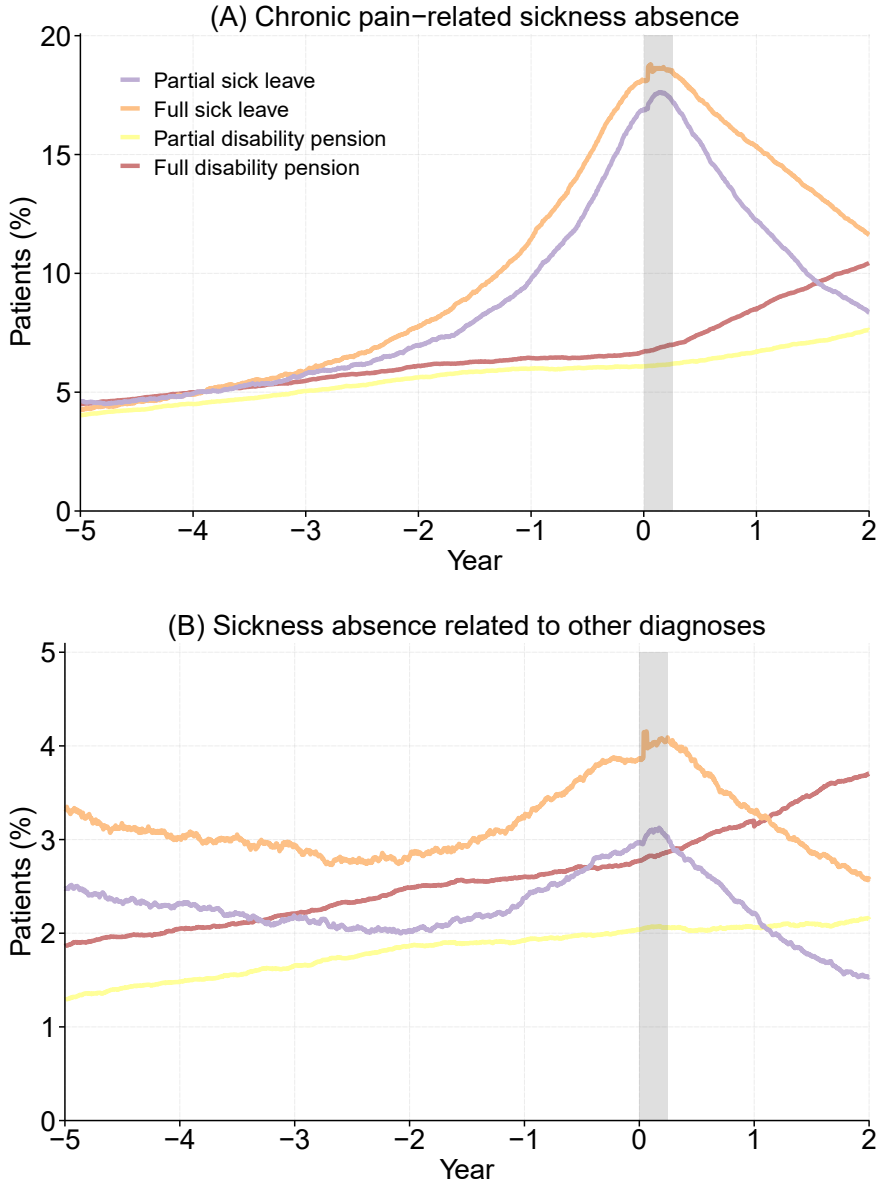

Figure 9. The daily sickness absence state distribution from five years before to two years after entry into specialist healthcare. Grey area denotes the most likely period in which a patient subsample was included in an IDT program. (A) Chronic pain-related sickness absence defined as ICD-10 codes: M(00−99), G(43−44, 47, 50−64, 82, 96−97), R(07, 10, 26, 29, 51−52), S(12−13, 22, 32, 42−43, 53), T(85, 88, 91−94), and F(32−33, 41, 43, 45). (B) Sickness absence related to non-chronic pain diagnoses (includes permanent disability pension related to chronic pain before 2005).

4.2 SICKNESS ABSENCE PREDICTION (STUDY I)

Patients were divided into three clusters based on their chronic pain-related SA in the two years following their first visit to a SQRP clinic (Figure 10). The first cluster contained patients with low SA that either received no social insurance benefits or recovered rapidly from such benefits (n=25,294). The second cluster was characterized by patients with high sick leave that either remained on sick leave until the end or stopped receiving sick leave benefits during the second year (n=5,416). The third cluster included patients with any amount of permanent disability pension (n=3,723). Finally, patients with permanent disability pension registered in the five years before entry into specialist healthcare were excluded from the prediction analyses to avoid inflated predictive performance (n=9,893).

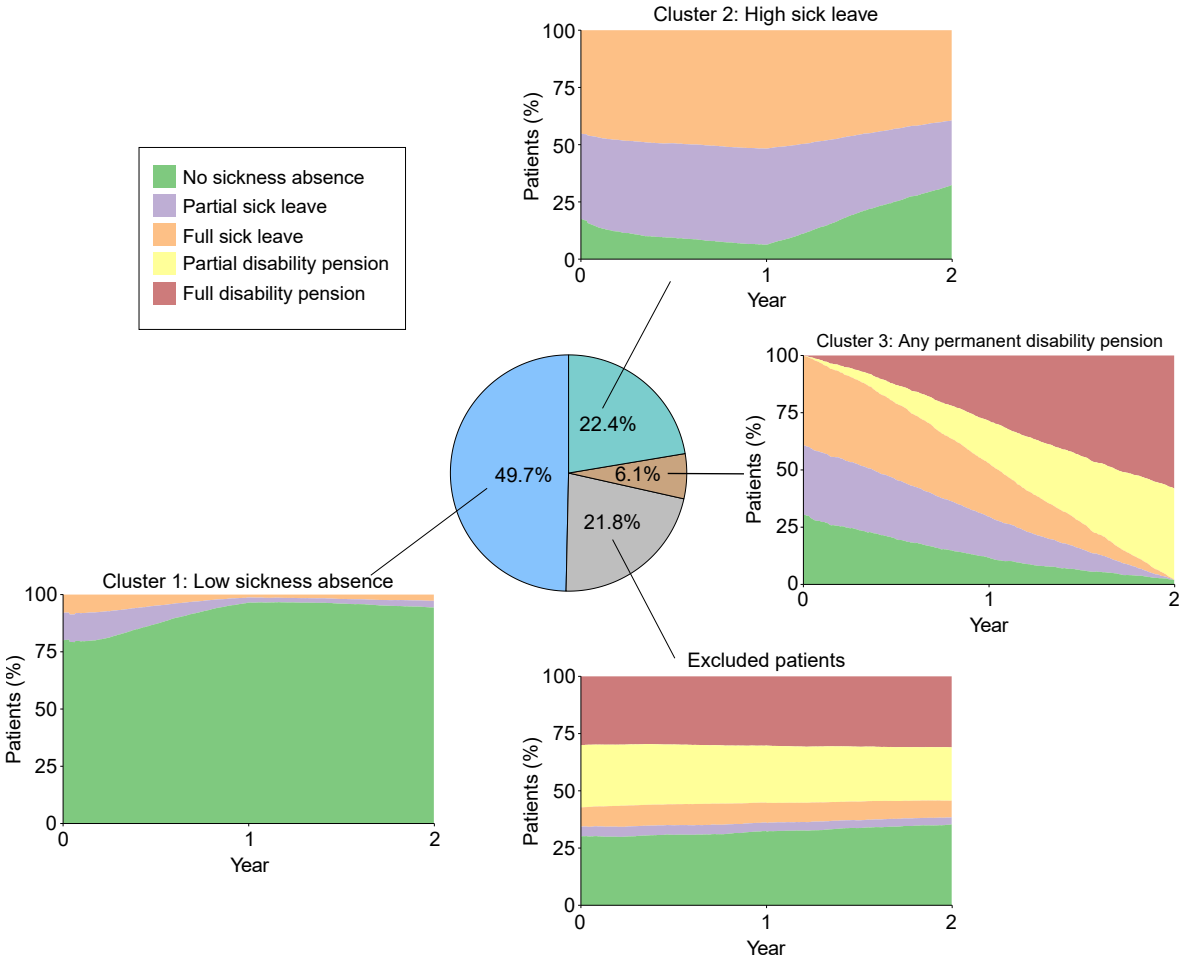

Figure 10. Cluster-wise daily distribution of chronic pain-related SA for the two years following the first visit to a SQRP clinic. The pie chart illustrates the patient distribution per cluster. Figure adopted from study II under CC BY 4.0.

In total, ten predictors of future SA were identified by the four algorithms (Figure 11). Information on sick leave history contributed the most to predictive performance and was included in the best models by all algorithms. Other consistently identified predictors were age and an indicator of a 2008 policy, while a southern Sweden geographical location and confidence in recovery were included by three of four algorithms. The gradient tree boosting

configuration was selected as the final model as it slightly but consistently outperformed the other algorithms. On average, it predicted 68% of the patients correctly across the three outcome classes compared to 33% by chance (balanced accuracy=0.68). The model was useful for predictions of low and high SA (sensitivity≥0.64; positive predictive value≥0.61), but not for future disability pensions as its low prevalence contributed to only one in four positive predictions being correct (sensitivity=0.70; positive predictive value=0.26). Instead, model performance was acceptable when collapsing high sick leave with disability pension (balanced accuracy=0.81; F1 score=0.79; sensitivity and specificity≥0.79; positive and negative predictive values≥0.70).

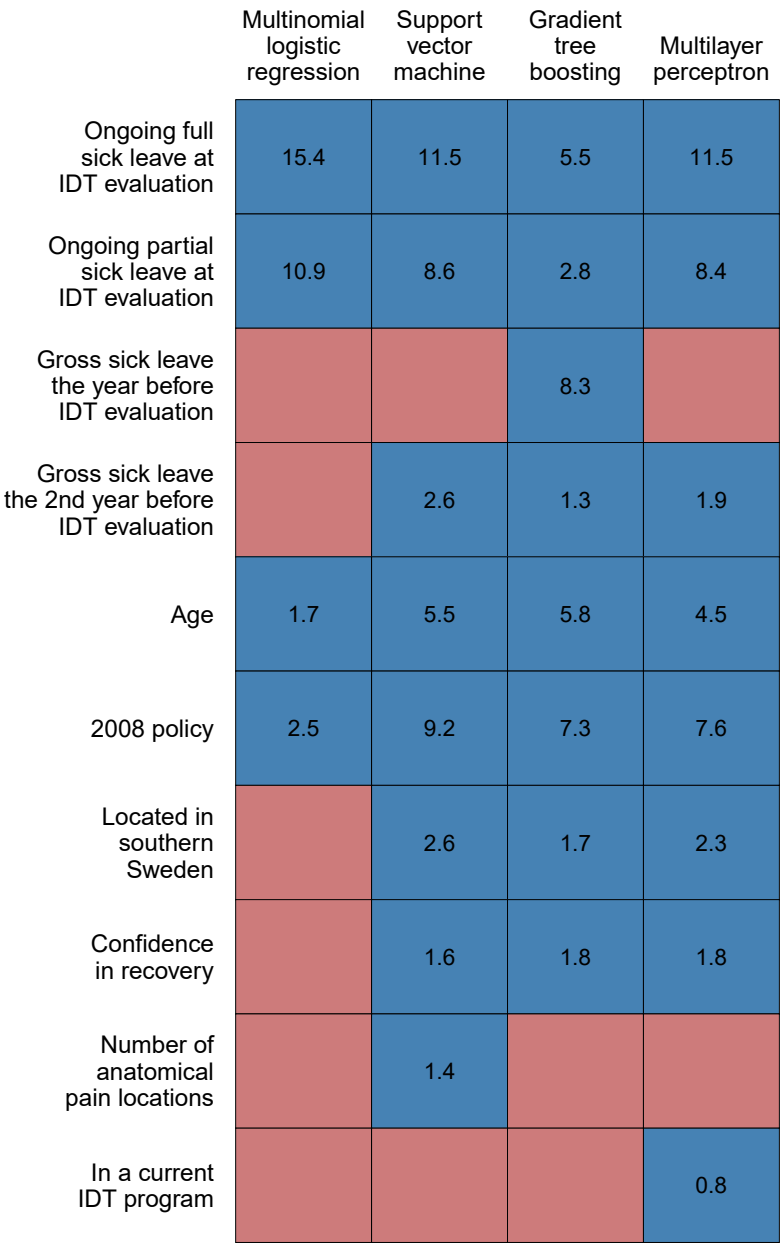

Figure 11. Predictors included in the final model configuration by algorithm. Blue and red denotes an included and non-included predictor, respectively. Numbers represent the loss in balanced accuracy for omitting the predictor.

4.3 SICKNESS ABSENCE INTERVENTION (STUDY II)

Of the 25,613 patients that met the eligibility criteria, 13,628 were IDT patients and 11,985 were patients assigned to other/no interventions. The intervention groups were similar in most measured baseline characteristics, apart from some variations in sex, geographical location, and employment status. SA trends were almost indistinguishable between groups from five years to one year before entry into specialist healthcare (Figure 13A). Sick leave then distinctly increased for IDT patients until the end of the assumed intervention period, rapidly decreased in the next year, and stayed slightly elevated over the remaining follow-up time compared to controls (Figure 13B). In contrast, disability pension trends were more similar between groups, with a steady increase from first visit to the clinic. SA patterns remained similar when stratifying on specialist clinic type (Figure 12). Over the five-year period, the mean difference in unadjusted and adjusted gross SA amounted to 75 days (95% CI: 58, 92) and 67 days (95% CI: 48, 87), respectively. Hence, neither observed nor population-average SA estimates supported any clinically meaningful effects in favor of IDT compared to other/no interventions.

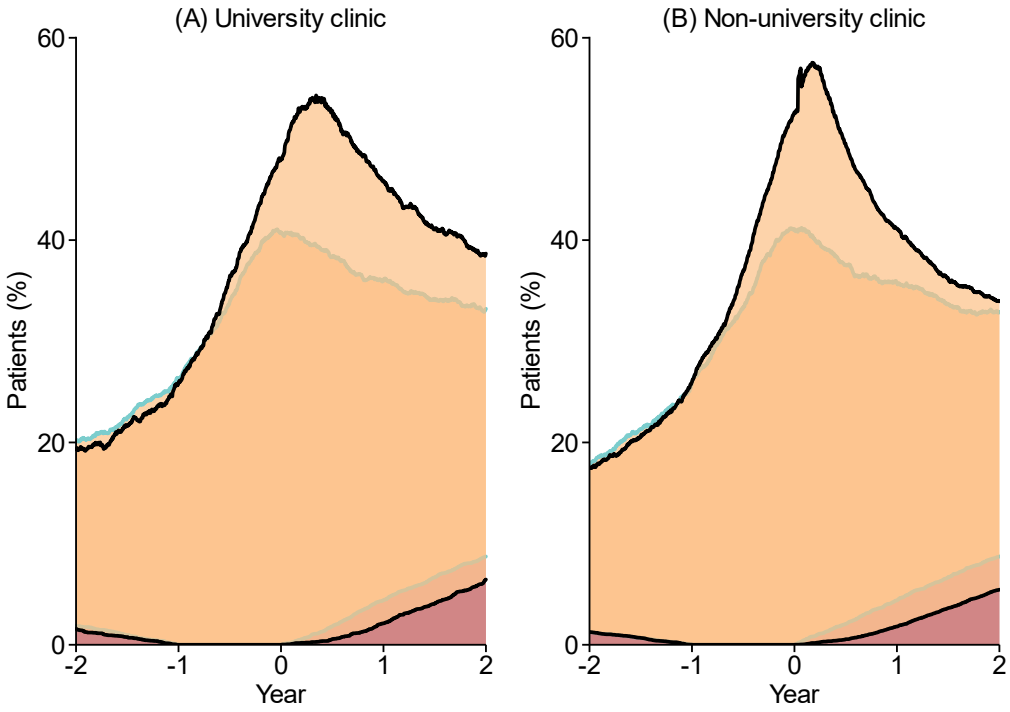

Figure 12. Unadjusted sickness absence by specialist clinic type. Colored areas display proportion of patients with sick leave (orange) and disability pension (red). Black and turquoise edges indicate IDT patients and patients in other/no interventions.

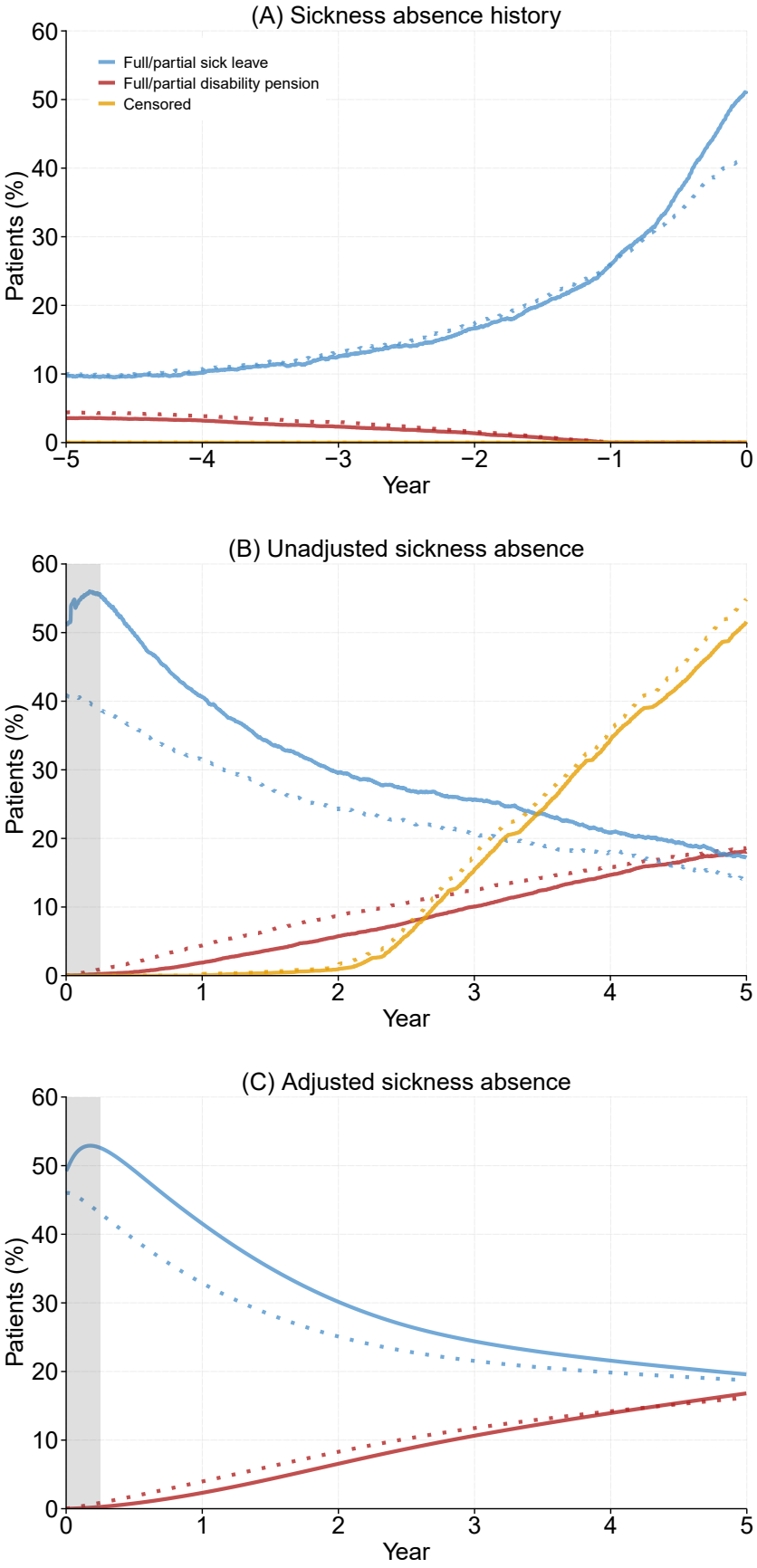

Figure 13. Sickness absence for patients included in an IDT program (solid line) and patients assigned to other/no interventions (dotted line). Time zero marks the first visit to the specialist clinic and grey areas denote the most likely period in which IDT programs took place.

4.4 PSYCHOMETRIC PROPERTIES OF CHRONIC PAIN EXPERIENCE QUESTIONNAIRES (STUDY III)

Figure 14 illustrates the empirical representations with the best properties under the IRT framework for all three questionnaires. SF-36 was most suitable for capturing the two independent traits of physical and mental health that were jointly estimated with subscale-specific residual factors. EQ-5D could be used as a unidimensional measure of HRQoL, while HADS best measured overall emotional distress when adjusted for the two residual factors of anxiety and depression. All three questionnaires were structurally sound (RMSEA ≤ 0.048; SRMSR ≤ 0.038) and their convergent/discriminant validity was supported by a logical association between the latent traits (Figure 15). However, whereas precision was acceptable for SF-36 and HADS (internal consistency reliability ≥ 0.79), it was inadequate for EQ-5D under the unidimensional IRT model (internal consistency reliability = 0.60).

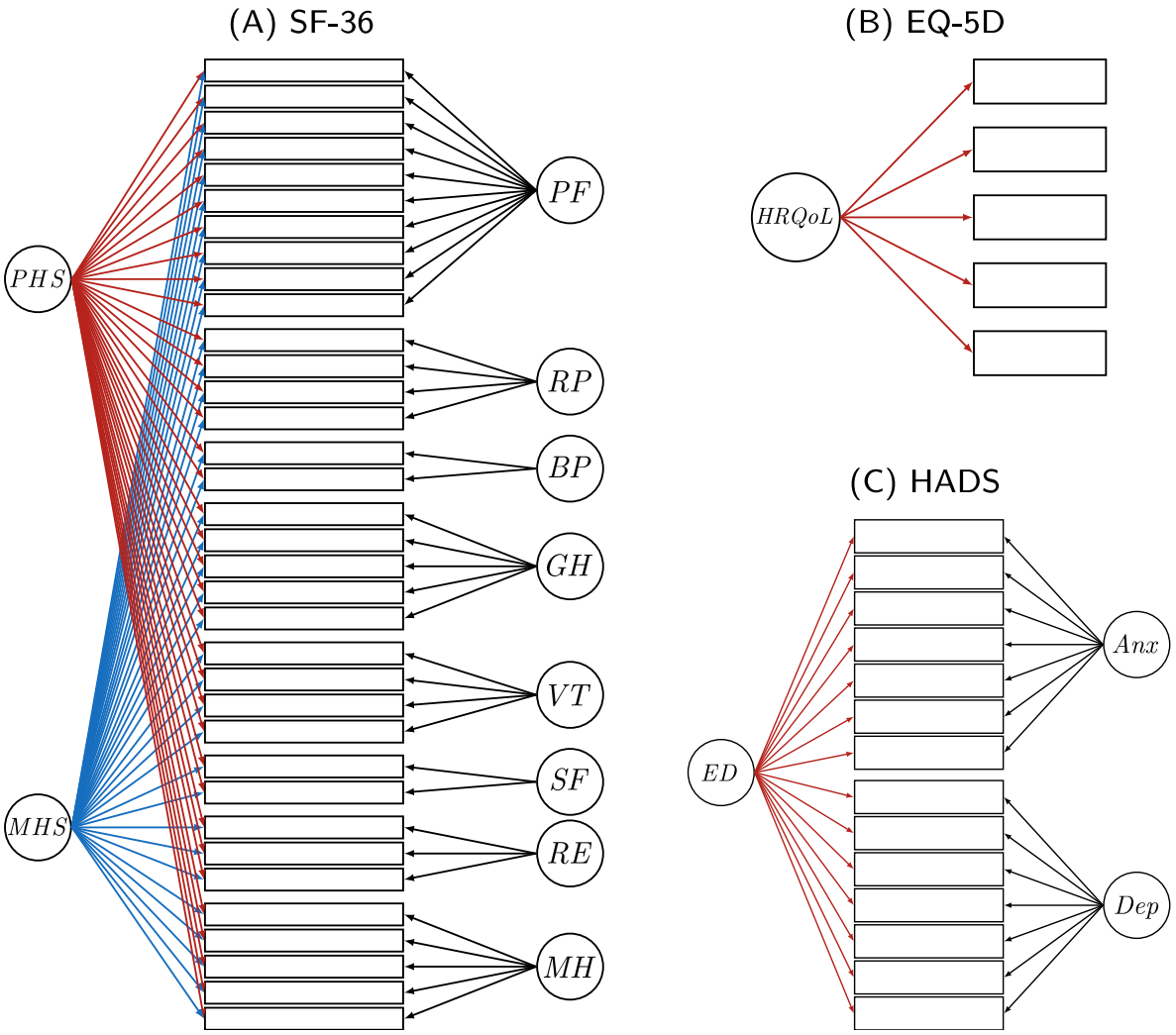

Figure 14. The empirical representations of each questionnaire with the best properties for chronic pain patients. Latent variables, assumed causal pathways, and questionnaire items are denoted by circles, arrows, and rectangles, respectively. PHS and MHS, physical and mental health summaries, respectively. PF, physical function. RP, role-physical limitations. BP, bodily pain. GH, general health. VT, vitality. SF, social function. RE, role-emotional limitations. MH, mental health. HRQoL, health-related quality of life. ED, emotional distress. Anx, anxiety. Dep, depression. Figures reproduced from study III under CC BY 4.0.

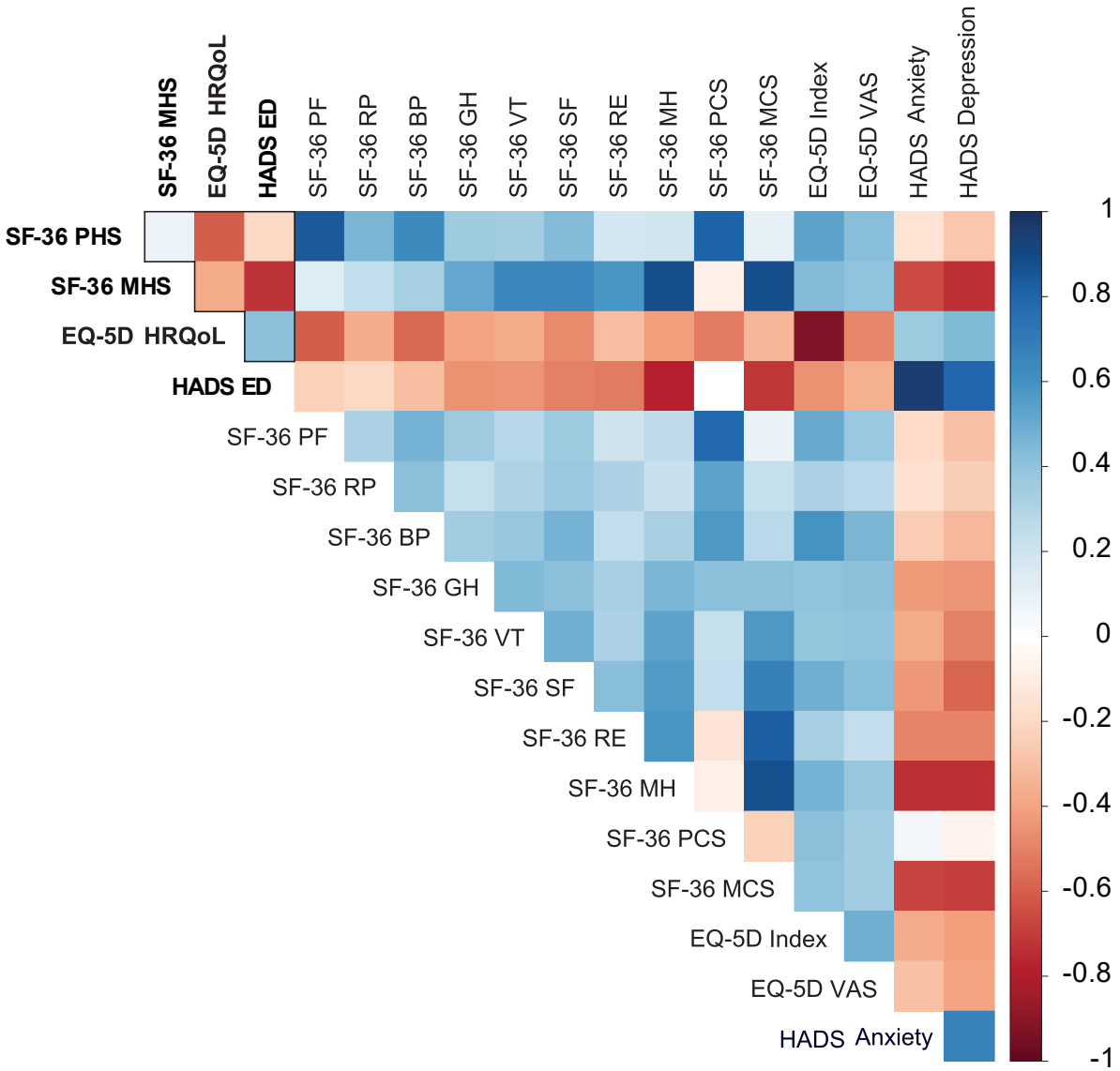

Figure 15. Correlation between IRT-based trait scores (bold) and conventionally estimated scores. PHS and MHS, physical and mental health summaries, respectively. HRQoL, health-related quality of life. ED, emotional distress. PF, physical functioning. RP, role-physical limitations. BP, bodily pain. GH, general health. VT, vitality. SF, social functioning. RE, role-emotional limitations. MH, mental health. PCS and MCS, physical and mental component summary scores, respectively. VAS, visual analogue scale. Figure reproduced from study III under CC BY 4.0.

5 DISCUSSION

5.1 MAIN FINDINGS

This thesis provided a SA overview, explored the possibility of SA prevention, and evaluated IDT as a SA intervention among chronic pain patients in Swedish specialist healthcare. It also assessed the psychometric properties of the instruments SF-36, EQ-5D, and HADS as core measures of the chronic pain experience.

Our results revealed a clear increase in SA over the five years before specialist healthcare entry, which then reversed and decreased over the next two years. Overall, nearly a fifth of the patients received SA benefits at any given time and almost half of them were on SA when they entered specialist healthcare. SA was also unevenly distributed so that a fifth of the patients accounted for nearly three-fifths of the SA net days. A model was developed with eight predictors of future SA that could differentiate between low and high SA at 80% accuracy. SA history contributed the most to predictive performance, while a temporal 2008-policy indicator, age, confidence in recovery, and geographical region were other predictors of lesser importance. Compared to unspecified treatments or recommendations, IDT did not offer any advantage in terms of SA reduction over a five-year period, under the assumption that the two groups were equivalent given our theory-based causal structure. Finally, our psychometric evaluation revealed that structurally sound and logically associated representations could be identified for all three questionnaires, but that EQ-5D was insufficiently precise as a unidimensional measure of HRQoL.

Sickness absence overview

Chronic pain disorders are recognized as leading contributors of prolonged SA, but SA overviews specific to chronic pain are rare due to difficulties in isolating the condition directly.[4,78,109,188] SQRP allowed us to describe SA in a well-defined clinical population with chronic pain impairments ample enough to motivate a specialist healthcare referral. Our findings are interesting from several perspectives. First, they provide a precise measure of the SA extent, which supports that SA is high among these patients. In the seven years surrounding specialist healthcare entry, patient-mean annual net SA ranged between 54 to 145 days compared to a maximum of 43 days for the general Swedish population over the 2000-2018 calendar period (Figure 16).[175] Second, they highlight the two years before to two years after specialist healthcare entry as a critical period for SA prevention and evaluation. This implicates that efforts to identify patients at the start of the period could prove to be valuable to mitigate the rapid SA increase in the coming two years, while intervention evaluations would capture the largest effect over the two years following IDT. Third, in agreement with previous reports, they show that SA is also unevenly distributed in this clinical population and that distinct SA strata can be isolated.[74,75,80,106,114] This implicates that large socioeconomic gains are possible by identifying and directing intensive intervention at the minority with the most SA.

Our findings also carry methodologically important implications. Because chronic pain-related SA is a strong indicator for IDT, the SA peak observed around specialist healthcare entry is likely a consequence of the patient selection procedure in the earlier healthcare chain.[15,61] The IDT referral indicates that patients are approaching their worse during this time and subsequent SA decrease could represent treatment response, natural course, placebo, Hawthorne effects, and statistical phenomena such as regression to the mean.[51,161] Moreover, the sharp SA increase following the first clinic visit likely reflects sick listing for IDT patients to enable program participation.[72] These two factors combined suggest that uncontrolled studies of interventions in specialist healthcare should be avoided as they artificially inflate treatment effects.

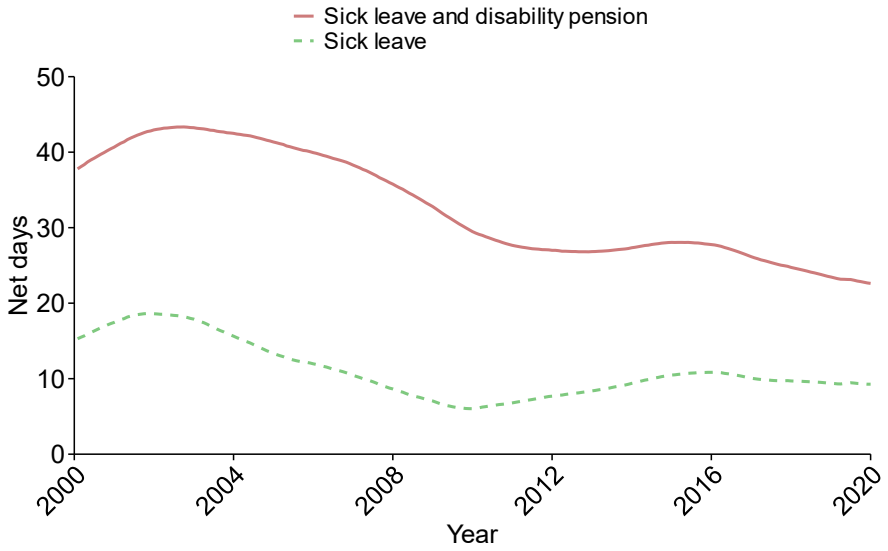

Figure 16. Sickness absence in the Swedish general population. Lines denote the moving 12-month mean for insured individuals aged 16-64. Source: Swedish Social Insurance Agency.

Sickness absence prediction

Early identification of high-risk patients is an important step in SA prevention. We examined whether future SA could be predicted from data known at the time of specialist healthcare entry and found that it was possible to discriminate between patients with future low and high SA at a meaningful accuracy. However, as interventions in practice are guided by clinical expertise, it would be more interesting to compare model predictions to the clinicians' expectations upon meeting the patient in the future.

In agreement with previous studies, SA history contained the most important information for identifying patients with future high sick leave or disability pension.[42,92,98,100,156,191] Performance was specific to chronic pain-related SA, which is consistent with earlier reports.[42,92] Two other clinically relevant predictors of future SA were age and confidence in recovery. In line with previous studies, our results suggested that the former was positively associated with future SA, while the latter was inversely associated with future disability

pension.[98,102,156,191] Meanwhile, two identified predictors of limited clinical importance were a 2008-policy indicator and geographical region, which aligned with previous reports of municipal and temporal SA differences.[103,108,167,191] Earlier SA risk factors that did not improve performance were sex, education, employment status, and multimorbidity.[98,156,191]

Sickness absence intervention

Despite its strong theoretical foundation and international status as a chronic pain core intervention, support for IDT is surprisingly limited.[44,97,129,160,186] The latest Cochrane evidence summary from 2015 revealed that the odds of return-to-work is the same for patients undergoing IDT as for those receiving usual care.[97] Other evidence summaries also generally agree that the support for IDT as a SA mitigator is equivocal.[44,97,160,186]

IDT in specialist healthcare is the most complex chronic pain intervention in Sweden today. Within this context, we conducted a pragmatic register-based study of its effects over a five-year period and found no support for IDT decreasing SA to a higher extent than alternative interventions. Our findings were consistent with three of seven peer-reviewed controlled studies published in the past decade that compared IDT to usual care over at least one year.[22,122,137] Two were randomized controlled trials that reported no differences in SA and return-to-work and one was a matched-cohort study that reported a slight advantage in SA for the controls. In addition, our results mirror the findings of a governmental evaluation in the form of a matched-cohort study of IDT in primary healthcare.[72,73] Conversely, four randomized controlled studies provided some support for IDT being superior to usual care.[12,23,79,155] Two of these studies reported IDT having up to large effects on SA, while the other two reported up to moderate effects on return to work. In summary, the evidence of IDT effects on SA remains inconclusive.

Our choice to use the patients merely assessed and offered other interventions or recommendations as controls was criticized in a commentary.[154] We agree that it introduces susceptibility to bias, as treatment assignment is influenced by clinical expertise and patient preference. Nevertheless, the intervention groups were similar in most measured baseline variables and our design strived towards further increasing their comparability by adjusting for theory-driven confounders identified in the scientific literature. As discussed in our reply, the proposed alternative of using the IDT patients as their own controls is not a viable solution, as the effects specific to IDT cannot be isolated in such designs.[110]

Psychometric properties of chronic pain experience questionnaires

SF-36, EQ-5D, and HADS are questionnaires widely implemented for measuring the chronic pain experience, but their properties are not sufficiently evaluated in the chronic pain population.[26,30,33,48] SQRP presented an excellent opportunity for assessing their

construct validity and internal consistency reliability under the psychometric IRT framework.

In agreement with its original theory, SF-36 was a valid and precise measure of two meta-constructs of physical and mental health.[77,193] Other studies typically also support these two constructs irrespective of the population investigated.[11,52,68,77,184,192,193] However, since the inception of SF-36, there is an ongoing debate on whether they should be perceived as independent or correlated.[77,193] Our results supported the former view, estimated in the form of a two-tier model, where all items loaded directly on both meta-constructs. To mitigate the known discord between the meta-constructs and domain-specific constructs, all item-trait loadings were constrained to positive values, but the problem was not completely eliminated and further improvements are likely possible.[177] Nevertheless, the physical and mental health constructs were logically associated both with the domain-specific constructs and with the constructs of EQ-5D and HADS.

Despite its underlying theory as a formative multidimensional scale, our results supported that EQ-5D was structurally valid, albeit imprecise, under the psychometric unidimensional model.[20] This is consistent with previous studies that also evaluated EQ-5D as a unidimensional scale in patients with chronic pain and mental illness.[93,142,166] However, because its theoretical foundation places each of the five items in a different dimension, we cannot exclude that the perceived unidimensionality was a consequence of insufficient sensitivity in fit evaluation methods.[34] It is nevertheless peculiar that the IRT-based HRQoL score was so strongly associated with the conventional index, despite their conceptual differences. However, this contradiction is of little practical concern as we discourage the use of EQ-5D as a unidimensional measure due to its low precision.[16]

HADS had the best properties as a measure of overall emotional distress, which conceptually closely relates to the anxiety and depression constructs defined in its original theory.[198] Whereas the bifactor model provided a valid and precise measure of emotional distress, HADS structural properties were not acceptable as a measure of anxiety and depression, which explains the criticism it has received in the past.[35] However, simply summarizing the score is not recommended, as the unidimensional HADS model had the worse properties of all structural models. Instead, the score should be adjusted for residual item dependencies of the anxiety and depression constructs. Our results are consistent with an earlier meta-study of mixed populations and largely also with the findings of two other studies that shared one chronic pain patient sample.[130,133,134] The combined research thereby supports that HADS is best used as a measure of overall emotional distress.

5.2 METHODOLOGICAL CONSIDERATIONS

Main limitations

It is critical to interpret scientific findings in light of their methodological limitations. Predictive studies are influenced by procedural decisions and can become obsolete if their underlying circumstances change, while systematic bias is the main threat to the internal validity of large-sample observational studies that estimate intervention effects, and psychometric evaluations are contingent on the questionnaire's underlying theory.

In study I, both the data management and the analyses required several methodological compromises. Numerous decisions were taken in the preparation of the SA data. First, the SA extent of each spell was treated as uniformly distributed, despite the knowledge that spells often decreased, increased, and varied over time. Second, non-permanent disability pension was reclassified as sick leave to better distinguish between permanent and non-permanent absenteeism, which was suboptimal and may complicate comparison with other studies. Finally, we differentiated between SA directly related to the chronic pain spectrum and other diagnoses for a more specific overview of its socioeconomic consequences.[70] However, a minimal but distinct SA increase around specialist healthcare entry indicated that other diagnoses were also influenced by the chronic pain condition (Figure 9B). With the SA data processed, sequence analysis was chosen because it provides a holistic overview of the SA trajectories and detailed insight into individual patient patterns.[152] This technique further makes no assumptions about the data-generating mechanisms, which aligned well with the exploratory approach taken in this study.[140] Our modelling procedure also had a direct impact on the interpretation of our results. Studies that explore the link between features and an outcome ultimately prioritize to either understand their relationship or to predict the outcome of individual analysis units.[17,162] Under the right circumstances, the former provides an accurate description of theoretical associations, but does not necessarily imply predictive ability.[162] Conversely, the latter excels in prediction at the cost of decreased accuracy and interpretability of the associations.[162] We took a machine learning-based approach, as our primary interest was to examine the extent to which future SA could be predicted at entry into specialist healthcare. The model with the highest performance did not provide a parsimonious summary of the associations between features and outcome, which instead needed to be interpreted indirectly. However, the many collinear features included among the predictor candidates, render it unlikely that a strictly regression-based approach would have been more informative.[17] As in any optimization procedure, there is no guarantee that the optimal model is identified unless all possible combinations are tested. Moreover, only a small portion of the hyperparameter space could be feasibly searched in the algorithm tuning. However, the parallel trial of four complementary algorithms resulted in models that largely confirmed each other. Noteworthy is that there generally was a considerable information redundancy within the feature domains and whereas the selected features represented those most important, other combinations approached a similar performance.

In study II, we estimated the population-average treatment effect of IDT assignment on SA under the framework of counterfactual causality. Target trial emulation is a strategy used for observational data to achieve the design-based inference of a randomized trial.[81,86,89,158] However, in addition to the statistical assumption of no model misspecification and no measurement error, causal inference necessitates the three untestable identifiability assumptions of consistency, positivity, and exchangeability.[81] First, the assumption of consistency asserts that observed and potentially factual outcomes are identical under a given circumstance, which implies that the intervention is sufficiently well-defined for consistent treatment effects. In our study, this assumption was likely violated due to the large heterogeneity of both IDT and control interventions across the specialist clinics. Nonetheless, we are confident that IDT systematically represented a higher level of engagement with the patient, given that it is the most extensive treatment offered. Second, the assumption of positivity proclaims that all the patients have a possibility of entering either intervention. Our data supported that this assumption was met, since both the distribution of measured baseline characteristics and examined covariate pattern frequencies were similar between the intervention groups. Third, the assumption of exchangeability affirms that the intervention groups are probabilistically equivalent at baseline. In our study, exchangeability was assumed conditionally on our causal structure (Figure 8) and subject to possible confounding bias, measurement bias, and selection bias.[81,104,157] Confounding bias is introduced via causes common to the intervention and outcome and is a major concern in observational studies, especially here since IDT was assigned in a clinical decision process. To manage this issue, we adjusted for important theory-driven confounders through inverse probability weights; however, some residual confounding likely remained, which can have biased our estimates in an unknown direction. This was possible either via suspected confounders that were unavailable to us or via other unknown confounders. Meanwhile, selection bias is introduced by conditioning on a common effect of both the intervention and outcome.[81] In our study, missingness can have influenced the results by opening non-causal pathways through conditioning on the patients with complete baseline data (Figure 17).[81,82] It would therefore have been preferable to manage the missingness rather than simply summarizing it. However, because the missingness was in the confounders, it could not simply be adjusted for in the inverse probability weights. Moreover, there was a limited selection of suitable auxiliary variables needed for multiple imputation, which further would have been rather computationally demanding when combined with the bootstrap confidence intervals recommended for inverse probability weighted-survival models.[9] Finally, measurement error cannot be described in terms of one general causal structure. Instead, it takes on several forms that are categorized under either non-differential or differential mismeasurement, where the errors of the intervention and outcome are independent and dependent, respectively.[81] Non-differential mismeasurement of intervention or outcome typically attenuates the causal effects through dilution, but imprecise measurement of confounders can also exacerbate them, since it results in them being incompletely partialled out.[81,157] In our data, the absence of information on SA spells below 15 days, retirement, emigration, and death introduced SA misclassification, which implies that the SA difference between intervention groups was

underestimated unless the SA misclassification was differential. This is, however, unlikely since it was prospectively measured by the external government apparatus. In contrast, mismeasurement of the self-report confounders emotional distress, everyday interference, and confidence in recovery likely biased the results in an unknown direction. Finally, besides the identifiability assumptions, model misspecification was possible though the logistic regression of the inverse probability weights, the flexible parametric survival models, and the Markov multistate model; model diagnostics were used to mitigate the possibility.[181] In summary, our results are contingent on several strong assumptions of which at least some were likely violated, thereby rendering it uncertain to what extent our estimates are reasonable approximations of the causal effect. However, given the magnitude of the negative effects found for IDT, considerable bias would be needed to obtain a clinically meaningful effect in favor of IDT.

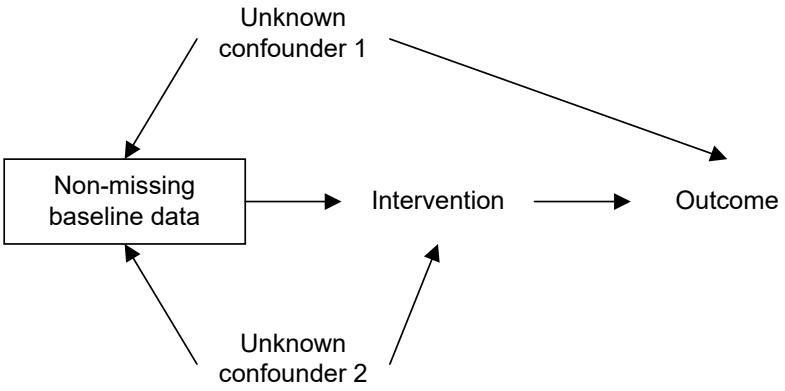

Figure 17. Directed acyclic graph of selection bias due to conditioning on patients with complete baseline data.

In study III, we evaluated the empirical properties of SF-36, EQ-5D, and HADS. The procedure relied upon indirect quality metrics contingent on the conceptual frameworks of the questionnaires, with cross-validation used to confirm the adequacy of the functional item-trait models.[149,151,162].[162] Nonetheless, the absence of an observed outcome makes the assessment of unsupervised problems complicated, which presented some practical difficulties. Most obviously, the conceptual model of EQ-5D as a multidimensional HRQoL measure with only one item per dimension prevented us from determining whether it was appropriately represented in a reflective model.[34] However, the strong relationship between the IRT-based reflective score and the traditional EQ-5D index nonetheless suggested that it ordered respondents on the HRQoL continuum in a similar manner, as previously reported.[143] In addition, our investigation was restricted to the marginal questionnaire properties, thereby assuming a uniform behavior in item responses irrespective of patient characteristics. It is possible that patients at the same latent trait status have tendencies to respond differently to individual items depending on their characteristics, which could invalidate our results if the population characteristics change markedly.[87]

Generalizability

Generalizability is dependent on eligibility criteria and sampling procedure. Our target population was chronic pain patients in Swedish specialist healthcare, while the source population was defined by SQRP, which consecutively aggregates patient data on a voluntary basis. A selection procedure on multiple levels could thus have decreased the sample representativeness relative to the target population.[104] First, the coverage of specialist clinics in SQRP was incomplete. This sample restriction was likely negligible, as 80-95% of clinics were included during the study period and non-participating clinics most probably had a low patient flow. Second, patients could decline to participate due to the voluntary consent requirement. Because there was no information on the decline rate and participation could be driven by patient characteristics, this selection represented the largest risk to generalizability. Finally, missing data on our study eligibility criteria potentially contributed to sample restriction if it systematically related to patient characteristics, but was likely negligible given the relatively small amounts of missingness. In light of these factors, the exact population characteristics remain unknown, leaving the possibility that a patient subgroup was not included in the study sample. Nevertheless, our results are relevant to adults with considerable chronic pain-related everyday impairments and a decreased HRQoL in Sweden.

The SA trends described in study I will likely remain similar, peaking around entry into specialist healthcare, albeit varying in absolute numbers. Likewise, the identified SA predictors were congruent with the results of several previous studies and are therefore likely to remain important in the future. However, even for the features evaluated here, the model composition and performance could differ with changed circumstances, as measurement error reduced performance and may have prevented otherwise important features from being selected in their current form. With respect to IDT, even when assuming adequate internal validity, the effects on SA reported in study II are highly dependent on program content and could change with the IDT development. Finally, the questionnaire measurement properties evaluated in study III should be generalizable to Swedish speakers forward in time unless the population characteristics change drastically. It is uncertain as to what extent they are applicable to other languages, but it is reassuring that our results corresponded well to previous studies of other translations.

6 CONCLUSIONS

This thesis has provided a sickness absence overview, explored the possibilities of sickness absence prevention, and evaluated interdisciplinary treatment as a sickness absence intervention among chronic pain patients in Swedish specialist healthcare. In addition, it presented psychometric information on three common questionnaires of the chronic pain experience.

- Sickness absence was high in the studied patients over the entire observation period. Temporally, it peaked around specialist healthcare entry, which was likely a consequence of the IDT referral procedure that prioritizes worsened patient status. Uncontrolled studies that simply attribute the decrease in sickness absence from the first visit to the clinic to an intervention are therefore prone to overestimating treatment effects.

- A data-driven model was developed that predicted at 80% accuracy whether patients would have low or high sickness absence in the coming two years. This suggests that it may be possible to identify patients that will have high sickness absence in the future already at baseline. Predictors of direct relevance to clinical practice included sick leave in the two preceding years, ongoing sick leave at entry into specialist healthcare, age, and confidence in recovery. Other predictors that were less informative were geographical location and a 2008-policy indicator, which nonetheless emphasized the importance of including spatial and temporal indicators in future predictive models.

- Our results showed no support for interdisciplinary treatment decreasing sickness absence compared to other/no interventions. Sickness absence trends were similar in both groups, albeit with interdisciplinary treatment patients receiving more social insurance benefits over a five-year period than other patients. Further research is needed to elucidate whether the results were a consequence of our methodology or represented the actual treatment effect. Given the inconclusive state of evidence, it nonetheless brings into question whether current programs are suitable for mitigating sickness absence.

- SF-36 and HADS are structurally and logically valid questionnaires with adequate precision for measuring core domains of the chronic pain experience. The former targets two independent traits of physical and mental health, while the latter is most suitable as an overall measure of emotional distress. Conversely, EQ-5D is not recommended as a unidimensional measure of HRQoL due to its insufficient precision.

7 FUTURE DIRECTIONS

Holistic life course perspective on chronic pain

We have described SA over a brief time period for chronic pain patients in specialist healthcare. With the diagnostic codes of ICD-11, new possibilities have presented themselves for identifying chronic pain in the general population, mapping their SA, and isolating its causes. Considering that chronic pain consequences extend to the social surrounding of the affected, it is also relevant to investigate concurrent and multigenerational consequences in their family for a more holistic view of the condition.

Decision support tools in clinical practice

Patients at risk of high future SA need to be identified earlier in the healthcare chain. Large socioeconomic gains will be made possible by preventing the SA increase two years before entry into specialist healthcare, which implicates identifying patients already in primary healthcare. Personalized medicine offers great advantages in other healthcare areas and decision support tools in clinical practice could prove valuable to optimize resource allocation for chronic pain patients.[94,168] Our prediction model only provides a crude measure of the possibilities in SA predictions and there is no reason to believe that performance could not be improved. With the ever-increasing data volumes, organisation and quality are important limiting factors, which emphasize the need for proper structure and adequate feature engineering.

Interdisciplinary treatment

Evidence is needed for IDT as a sickness absence mitigator. Despite its theoretical appeal and international recognition as a chronic pain core intervention, scientific evidence of its effects on SA is surprisingly limited. In part, this is understandable given the combined complexity of chronic pain and IDT; however, it emphasizes the need for more rigorous studies. Properly designed and conducted randomized controlled trials remain the cornerstone of causal inference, which in combination with the Swedish National Registers would permit cost-effective long-term follow-up with minimal attrition.[81] These registers are also important for pragmatic observational trials to establish generalizability to real-world practice. If primary healthcare data is incorporated into the National Patient Register, as proposed by the National Board of Health and Welfare, several of the limitations of National Quality Registers could be overcome.[126] Finally, it is important to discourage the practice of uncontrolled before-and-after studies of IDT effects on SA given their severe limitations and minimal contribution to the state of evidence.[38,51,161]

Consensus on chronic pain experience questionnaires

A broader consensus on core domain questionnaires of the chronic pain experience is needed. Few established instruments would lay the foundation for a better understanding of the condition, facilitate comparison, and increase coherence of intervention evaluation. A central authority, such as IASP, could be tasked with issuing recommendations aiming to improve the likelihood of widespread implementation. Selected questionnaires need to be both theoretically motivated and psychometrically sound. Here, we evaluated some measurement properties of three previously recommended generic questionnaires, but even for these instruments, many properties remain to be examined in chronic pain patients, including content validity, retest reliability, and responsiveness to change.[30]

Swedish Quality Register for Pain Rehabilitation

SQRP is a valuable complement to the Swedish centrally governed registers for providing insight into the chronic pain experience. Unfortunately, it has several limitations that effectively restrict its relevance to clinical research. Given that the purpose of quality registers is to monitor and improve healthcare quality and equality, it is of public interest that these limitations are resolved.[170] The following modifications are proposed for augmenting data quality to an acceptable level for pragmatic IDT evaluation. First, a better overview of SQRP patients is needed to assess how selection affects internal validity and generalizability. A feasible solution would be to routinely collect information on specialist clinic referral rates, IDT admission rates, and SQRP acceptance rates, combined with aggregate data on patient characteristics. Second, SQRP variables should be updated to better cover the domains of the chronic pain experience. Current information is mostly acceptable for pain characteristics and emotional distress, while physical and social domains are either inappropriately measured or absent at the cost of variables that are readily available in other registers. Adequate variable selection is complicated and is probably best determined by interdisciplinary competences with the purpose of the quality register in mind. Third, more detailed information on IDT program characteristics is needed to assure their quality and improve treatment effect assessment. No such data is currently stored in SQRP besides whether patients were assigned to an IDT program. Important details to include are program duration, intervention modules, involved care personnel, and patient compliance. Fourth, there is an urgent need to identify an appropriate control group to evaluate the IDT effect. Studies that simply attribute natural course to IDT are prone to overestimating treatment effects.[38,51,161] Because it is both ethically and practically problematic to isolate valid control groups in the clinical setting, non-IDT SQRP patients should be further explored as a viable alternative, given their similarity to IDT patients in measured baseline variables. Information on clinical decision criteria is therefore necessary to understand the mechanism behind the selection procedure and how these patients may differ. In the best case, patients assigned to treatments other than IDT would also be followed-up after departure from specialist healthcare.

8 ACKNOWLEDGEMENTS

Many influences have contributed to this thesis.

Lea Constan, wife, for giving me the extra push when needed; it would not have been possible without your support and knowledge.

Björn Äng, main supervisor, for a great opportunity for self-development, support, trust, and tolerance in my decisions. You are the best juggler that I know and still always in a pleasant mood; a great role model.

Paolo Frumento, mentor, for your support and wisdom; your intelligence is superseded by your kindness only.

Linda Vixner, co-supervisor, for your unusual stability and reliability; the key to any door.

Elena Tseli, **Tony Bohman**, **Andreas Monnier**, **Veronica Sjöberg**, and **Jens Westergren**, HD colleagues, for all the good times.

Örjan Dahlström, Mathilda Björk, and **Björn Gerdle**, Linköping collaborators, thank you for your contribution, I have learned from our interactions.

My **KI research group** for interesting discussions and the thesis seminar. Specifically, **Wim Grooten** for always lightening up the surrounding with your positive attitude to life, **Eva Rasmussen-Barr** for our fruitful collaborations, and **Lena Nilsson-Wikmar** for discovering me.

Pahansen de Alwis and **Karl Garme**, KTH colleagues, for the insight into your fascinating world.

My colleagues at the **CKF Falun** for stimulating discussions, a pleasant work environment, and the pre-disputation seminar. Specifically, **Erica Schytt** for your kindness and understanding in the final period of the thesis writing.

Lars Rönnegård, HD professor, for the excellent feedback on my thesis.

Family and friends, for your great support.

The Open-Source Community for subverting the power; making the world a better place.

The doctoral school in Epidemiology for great courses and dedicated teachers.

The **KI administrative personnel** for all the help with everyday practicalities.

Finally, **everyone else** that has been involved.

Grants: Swedish Research Council (Vetenskapsrådet: 2015-02512) and Swedish Research Council for Health, Working Life and Welfare (FORTE: 2016-07414 and 2017-00177).

Research time: Center for Clinical Research Falun

There are no facts, only interpretations.

Friedrich Nietzsche

9 REFERENCES

1 Aalen, OO & Johansen, S. An empirical transition matrix for nonhomogeneous Markov chains based on censored observation. Scand J Stat. 1978;5:141–150.

2 Adler, NE & Newman, K. Socioeconomic Disparities In Health: Pathways And Policies. Health Aff (Millwood). 2002;21(2):60-76. doi:10.1377/hlthaff.21.2.60.

3 Alexanderson, K. Sickness absence: a review of performed studies with focused on levels of exposures and theories utilized. Scand J Soc Med. 1998;26(4):241-249. doi:10.1177/14034948980260040301.

4 Alexanderson, K & Norlund, A. Sickness absence: causes, consequences, and physicians' sickness certification practice. A systematic literature review by the Swedish Council on Technology Assessment in Health Care (SBU). Scand J Public Health Suppl. 2004;63:3-263.

5 Allebeck, P & Mastekaasa, A. Swedish Council on Technology Assessment in Health Care (SBU). Chapter 3. Causes of sickness absence: research approaches and explanatory models. Scand J Public Health Suppl. 2004;63:36-43. doi:10.1080/14034950410021835.

6 Andersen, LN et al. Psychosocial aspects of everyday life with chronic musculoskeletal pain: A systematic review. Scand J Pain. 2014;5(2):131-148. doi:10.1016/j.sjpain.2014.01.001.

7 Andersen, ML, Araujo, P, Frange, C & Tufik, S. Sleep Disturbance and Pain: A Tale of Two Common Problems. Chest. 2018;154(5):1249-1259. doi:10.1016/j.chest.2018.07.019.

8 Andersson, GB. Epidemiological features of chronic low-back pain. Lancet. 1999;354(9178):581-585. doi:10.1016/S0140-6736(99)01312-4.

9 Austin, PC. Variance estimation when using inverse probability of treatment weighting (IPTW) with survival analysis. Stat Med. 2016;35(30):5642-5655. doi:10.1002/sim.7084.

10 Ballabio, D, Grisoni, F & Todeschini, R. Multivariate comparison of classification performance measures. Chemometr Intell Lab. 2018;174:33-44.

11 Beals, J et al. Different factor loadings for SF36: the Strong Heart Study and the National Survey of Functional Health Status. J Clin Epidemiol. 2006;59(2):208-215. doi:10.1016/j.jclinepi.2005.07.010.

12 Berglund, E et al. Multidisciplinary Intervention and Acceptance and Commitment Therapy for Return-to-Work and Increased Employability among Patients with Mental Illness and/or Chronic Pain: A Randomized Controlled Trial. Int J Environ Res Public Health. 2018;15(11). doi:10.3390/ijerph15112424.

13 Bonica, J. The Management of Pain. Philadelphia, PA; Lea and Febirger, 1953.

14 Borsboom, D, Mellenbergh, GJ & van Heerden, J. The theoretical status of latent variables. Psychol Rev. 2003;110(2):203-219. doi:10.1037/0033-295X.110.2.203.

15 Bramberg, EB et al. Implementation of evidence-based rehabilitation for non-specific back pain and common mental health problems: a process evaluation of a nationwide initiative. BMC Health Serv Res. 2015;15:79. doi:10.1186/s12913-015-0740-4.

16 Brazier, J et al. A systematic review, psychometric analysis and qualitative assessment of generic preference-based measures of health in mental health populations and the estimation of mapping functions from widely used specific measures. Health Technol Assess. 2014;18(34):vii-viii, xiii-xxv, 1-188. doi:10.3310/hta18340.

17 Breiman, L. Statistical modeling: The two cultures. Statistical Science. 2011;16:199-231.

18 Breivik, H, Collett, B, Ventafridda, V, Cohen, R & Gallacher, D. Survey of chronic pain in Europe: prevalence, impact on daily life, and treatment. Eur J Pain. 2006;10(4):287-333. doi:10.1016/j.ejpain.2005.06.009.

19 Breivik, H, Eisenberg, E, O'Brien, T & Openminds. The individual and societal burden of chronic pain in Europe: the case for strategic prioritisation and action to improve knowledge and availability of appropriate care. BMC Public Health. 2013;13:1229. doi:10.1186/1471-2458-13-1229.

20 Brooks, R. EuroQol: the current state of play. Health Policy. 1996;37(1):53-72. doi:10.1016/0168-8510(96)00822-6.

21 Burke, AL, Mathias, JL & Denson, LA. Psychological functioning of people living with chronic pain: a meta-analytic review. Br J Clin Psychol. 2015;54(3):345-360. doi:10.1111/bjc.12078.

22 Busch, H, Bjork Bramberg, E, Hagberg, J, Bodin, L & Jensen, I. The effects of multimodal rehabilitation on pain-related sickness absence - an observational study. Disabil Rehabil. 2018;40(14):1646-1653. doi:10.1080/09638288.2017.1305456.

23 Busch, H, Bodin, L, Bergstrom, G & Jensen, IB. Patterns of sickness absence a decade after pain-related multidisciplinary rehabilitation. Pain. 2011;152(8):1727-1733. doi:10.1016/j.pain.2011.02.004.

24 Chalmers, RP. mirt: a multidimensional item response theory package for the R environment. J Stat Softw. 2012;48:1–29.

25 Chapman, CR, Tuckett, RP & Song, CW. Pain and stress in a systems perspective: reciprocal neural, endocrine, and immune interactions. J Pain. 2008;9(2):122-145. doi:10.1016/j.jpain.2007.09.006.

26 Chapman, JR et al. Evaluating common outcomes for measuring treatment success for chronic low back pain. Spine (Phila Pa 1976). 2011;36(21 Suppl):S54-68. doi:10.1097/BRS.0b013e31822ef74d.

27 Chavent, M, Lechevallier, Y & Briant, O. DIVCLUS-T: A monothetic divisive hierarchical clustering method. Comput Stat Data Anal. 2007;52:687-701.

28 Chen, T & Guestrin, G. XGBoost: A scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining. Association for Computing Machinery, 2016.

29 Chiarotto, A et al. Core outcome domains for clinical trials in non-specific low back pain. Eur Spine J. 2015;24(6):1127-1142. doi:10.1007/s00586-015-3892-3.

30 Chiarotto, A, Terwee, CB, Kamper, SJ, Boers, M & Ostelo, RW. Evidence on the measurement properties of health-related quality of life instruments is largely missing in patients with low back pain: A systematic review. J Clin Epidemiol. 2018;102:23-37. doi:10.1016/j.jclinepi.2018.05.006.

31 Clements, M et al. rstpm2: Smooth Survival Models, Including Generalized Survival Models. The Comprehensive R Archive Network, 2019. url: https://CRAN.R-project.org/package=rstpm2.

32 Cohen, SP, Vase, L & Hooten, WM. Chronic pain: an update on burden, best practices, and new advances. Lancet. 2021;397(10289):2082-2097. doi:10.1016/S0140-6736(21)00393-7.

33 Cosco, TD, Doyle, F, Ward, M & McGee, H. Latent structure of the Hospital Anxiety And Depression Scale: a 10-year systematic review. J Psychosom Res. 2012;72(3):180-184. doi:10.1016/j.jpsychores.2011.06.008.

34 Costa, DSJ. Is health-related quality of life a useful outcome in pain intervention research? Pain. 2020;161(1):1-2. doi:10.1097/j.pain.0000000000001701.

35 Coyne, JC & van Sonderen, E. No further research needed: abandoning the Hospital and Anxiety Depression Scale (HADS). J Psychosom Res. 2012;72(3):173-174. doi:10.1016/j.jpsychores.2011.12.003.

36 Craig, P et al. Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ. 2008;337:a1655. doi:10.1136/bmj.a1655.

37 Crowther, MJ & Lambert, PC. Parametric multistate survival models: Flexible modelling allowing transition-specific distributions with application to estimating clinically useful measures of effect differences. Stat Med. 2017;36(29):4719-4742. doi:10.1002/sim.7448.

38 Cuijpers, P, Weitz, E, Cristea, IA & Twisk, J. Pre-post effect sizes should be avoided in meta-analyses. Epidemiol Psychiatr Sci. 2017;26(4):364-368. doi:10.1017/S2045796016000809.

39 Cylus, J, Hartman, M, Washington, B, Andrews, K & Catlin, A. Pronounced gender and age differences are evident in personal health care spending per person. Health Aff (Millwood). 2011;30(1):153–160. doi:10.1377/hlthaff.2010.0216.

40 de Wreede, L, Fiocco, M & Putter, H. mstate: An R package for the analysis of competing risks and multi-state models. J Stat Softw. 2011;38(7):1–30.

41 Demyttenaere, K et al. Mental disorders among persons with chronic back or neck pain: results from the World Mental Health Surveys. Pain. 2007;129(3):332-342. doi:10.1016/j.pain.2007.01.022.

42 Dorner, TE et al. Sickness absence due to back pain or depressive episode and the risk of all-cause and diagnosis-specific disability pension: A Swedish cohort study of 4,823,069 individuals. Eur J Pain. 2015;19(9):1308-1320. doi:10.1002/ejp.661.

43 Dowle, M & Srinivasan, A. data.table: Extension of data.frame. The Comprehensive R Archive Network, 2020. url: https://CRAN.R-project.org/package=data.table.

44 Dragioti, E, Evangelou, E, Larsson, B & Gerdle, B. Effectiveness of multidisciplinary programmes for clinical pain conditions: An umbrella review. J Rehabil Med. 2018;50(9):779-791. doi:10.2340/16501977-2377.

45 Duenas, M, Ojeda, B, Salazar, A, Mico, JA & Failde, I. A review of chronic pain impact on patients, their social environment and the health care system. J Pain Res. 2016;9:457-467. doi:10.2147/JPR.S105892.

46 Edwards, JR. The fallacy of formative measurement. Organ Res Methods. 2011;14(4):370-388.

47 Edwards, JR & Bagozzi, RP. On the nature and direction of relationships between constructs and measures. Psychol Methods. 2000;5(2):155-174. doi:10.1037/1082-989x.5.2.155.

48 Edwards, RR et al. Patient phenotyping in clinical trials of chronic pain treatments: IMMPACT recommendations. Pain. 2016;157(9):1851-1871. doi:10.1097/j.pain.0000000000000602.

49 Elzahaf, RA, Tashani, OA, Unsworth, BA & Johnson, MI. The prevalence of chronic pain with an analysis of countries with a Human Development Index less than 0.9: a systematic review without meta-analysis. Curr Med Res Opin. 2012;28(7):1221-1229. doi:10.1185/03007995.2012.703132.

50 Engel, GL. The need for a new medical model: a challenge for biomedicine. Science. 1977;196(4286):129-136. doi:10.1126/science.847460.

51 Evans, SR. Clinical trial structures. J Exp Stroke Transl Med. 2010;3(1):8-18. doi:10.6030/1939-067x-3.1.8.

52 Farivar, SS, Cunningham, WE & Hays, RD. Correlated physical and mental health summary scores for the SF-36 and SF-12 Health Survey, V.I. Health Qual Life Outcomes. 2007;5:54. doi:10.1186/1477-7525-5-54.

53 Fayaz, A, Croft, P, Langford, RM, Donaldson, LJ & Jones, GT. Prevalence of chronic pain in the UK: a systematic review and meta-analysis of population studies. BMJ Open. 2016;6(6):e010364. doi:10.1136/bmjopen-2015-010364.

54 Fayers, PM. Causal variables, indicator variables and measurement scales: an example from quality of life. J R Statist Soc A. 2002;165:233-261.

55 Forsberg Stjernschantz, J. [Registerforskning – Etiken bakom juridiken]. Stockholm, Karolinska Institutet and Uppsala University, 2013. Report No: 978-91-981256-1-0.

56 Gabadinho, A, Ritschard, G, Müller, NS & Studer, M. Analyzing and visualizing state sequences in R with TraMineR. J Stat Softw. 2011;40:1-37.

57 Gaskin, DJ & Richard, P. The economic costs of pain in the United States. J Pain. 2012;13(8):715-724. doi:10.1016/j.jpain.2012.03.009.

58 Gatchel, RJ, McGeary, DD, McGeary, CA & Lippe, B. Interdisciplinary chronic pain management: past, present, and future. Am Psychol. 2014;69(2):119-130. doi:10.1037/a0035514.

59 Gatchel, RJ, Peng, YB, Peters, ML, Fuchs, PN & Turk, DC. The biopsychosocial approach to chronic pain: scientific advances and future directions. Psychol Bull. 2007;133(4):581-624. doi:10.1037/0033-2909.133.4.581.

60 GBD 2019 Diseases and Injuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet. 2020;396(10258):1204-1222. doi:10.1016/S0140-6736(20)30925-9.

61 Gerdle, B, Stalnacke, BM, Soderlund, A & Asenlöf, P. [Indikation för multimodal rehabilitering vid långvarig smärta (Rapport 2011:02)]. Samarbetsprojektet nationella medicinska indikationer, 2011.

62 Géron, A. Hands-on Machine Learning With Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems. Sebastopol, CA; O'Reilly Media, 2019.

63 Global Burden of Disease Collaborative Network. Global Burden of Disease Study 2019 (GBD 2019) Results. url: http://ghdx.healthdata.org/gbd-results-tool. Accessed: 2020.

64 Godoy, LD, Rossignoli, MT, Delfino-Pereira, P, Garcia-Cairasco, N & de Lima Umeoka, EH. A Comprehensive Overview on Stress Neurobiology: Basic Concepts and Clinical Implications. Front Behav Neurosci. 2018;12:127. doi:10.3389/fnbeh.2018.00127.

65 Goldberg, DS & McGee, SJ. Pain as a global public health priority. BMC Public Health. 2011;11:770. doi:10.1186/1471-2458-11-770.

66 Goldstein, DS. How does homeostasis happen? Integrative physiological, systems biological, and evolutionary perspectives. Am J Physiol Regul Integr Comp Physiol. 2019;316(4):R301-R317.

67 Gran, JM, Lie, SA, Oyeflaten, I, Borgan, O & Aalen, OO. Causal inference in multi-state models-sickness absence and work for 1145 participants after work rehabilitation. BMC Public Health. 2015;15:1082. doi:10.1186/s12889-015-2408-8.

68 Grassi, M, Nucera, A & European Community Respiratory Health Study Quality of Life Working, G. Dimensionality and summary measures of the SF-36 v1.6: comparison of scale- and item-based approach across ECRHS II adults population. Value Health. 2010;13(4):469-478. doi:10.1111/j.1524-4733.2009.00684.x.

69 Greenland, S, Pearl, J & Robins, JM. Causal diagrams for epidemiologic research. Epidemiology. 1999;10(1):37-48.

70 Gustavsson, A et al. Socio-economic burden of patients with a diagnosis related to chronic pain--register data of 840,000 Swedish patients. Eur J Pain. 2012;16(2):289-299. doi:10.1016/j.ejpain.2011.07.006.

71 Hagglund, P & Johansson, P. [Upp till bevis – en ESO-rapport om experiment som underlag för politik (Rapport: 2021:4)]. Stockholm, 2021. Report No: 978-91-525-0214-3.

72 Hagglund, P, Johansson, P & Laun, L. [Rehabiliteringsgarantin (Rapport 2012:26)]. Uppsala, Institute for evaluation of labour market and education policy, 2012.

73 Hagglund, P, Johansson, P & Laun, L. [Insatserna inom rehabiliteringsgarantin och deras effekter på hälsa och sjukfrånvaro (Rapport 2014:12)]. Uppsala, Institute for evaluation of labour market and education policy, 2014.

74 Hashemi, L, Webster, BS & Clancy, EA. Trends in disability duration and cost of workers' compensation low back pain claims (1988-1996). J Occup Environ Med. 1998;40(12):1110-1119. doi:10.1097/00043764-199812000-00011.

75 Hashemi, L, Webster, BS, Clancy, EA & Courtney, TK. Length of disability and cost of work-related musculoskeletal disorders of the upper extremity. J Occup Environ Med. 1998;40(3):261-269. doi:10.1097/00043764-199803000-00008.

76 Hastie, T, Tibshirani, R & Friedman, J. The Elements of Statistical Learning: Data mining, Inference, and Prediction. New York; Springer, 2017.

77 Hays, RD, Sherbourne, CD & Mazel, RM. The RAND 36-Item Health Survey 1.0. Health Econ. 1993;2(3):217-227. doi:10.1002/hec.4730020305.

78 Head, J et al. Diagnosis-specific sickness absence as a predictor of mortality: the Whitehall II prospective cohort study. BMJ. 2008;337:a1469. doi:10.1136/bmj.a1469.

79 Henchoz, Y, de Goumoens, P, So, AK & Paillex, R. Functional multidisciplinary rehabilitation versus outpatient physiotherapy for non specific low back pain: randomized controlled trial. Swiss Med Wkly. 2010;140:w13133. doi:10.4414/smw.2010.13133.

80 Henderson, M, Glozier, N & Holland Elliott, K. Long term sickness absence. BMJ. 2005;330(7495):802-803. doi:10.1136/bmj.330.7495.802.

81 Hernan, M & Robins, J. Causal inference: What if. Chapman & Hall/CRC Press, 2020.

82 Hernan, MA, Hernandez-Diaz, S & Robins, JM. A structural approach to selection bias. Epidemiology. 2004;15(5):615-625. doi:10.1097/01.ede.0000135174.63482.43.

83 Hernan, MA & Robins, JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. Am J Epidemiol. 2016;183(8):758-764. doi:10.1093/aje/kwv254.

84 Hernan, MA & Robins, JM. Per-Protocol Analyses of Pragmatic Trials. N Engl J Med. 2017;377(14):1391-1398. doi:10.1056/NEJMsm1605385.

85 Hestbaek, L, Leboeuf-Yde, C & Manniche, C. Low back pain: what is the long-term course? A review of studies of general patient populations. Eur Spine J. 2003;12(2):149-165. doi:10.1007/s00586-002-0508-5.

86 Holland, PW. Statistics and causal inference. J Am Stat Assoc. 1986;81(396):945–960.

87 Holland, PW & Howard, W. Differential item functioning. Hillsdale; Lawrence Erlbaum Associates, 1993.

88 Hooten, WM. Chronic Pain and Mental Health Disorders: Shared Neural Mechanisms, Epidemiology, and Treatment. Mayo Clin Proc. 2016;91(7):955-970. doi:10.1016/j.mayocp.2016.04.029.

89 Imbens, GW & Rubin, DB. Causal inference in statistics, social, and biomedical sciences. Cambridge University Press, 2015.

90 Itz, CJ, Geurts, JW, van Kleef, M & Nelemans, P. Clinical course of non-specific low back pain: a systematic review of prospective cohort studies set in primary care. Eur J Pain. 2013;17(1):5-15. doi:10.1002/j.1532-2149.2012.00170.x.

91 Jackson, T et al. A Systematic Review and Meta-Analysis of the Global Burden of Chronic Pain Without Clear Etiology in Low- and Middle-Income Countries: Trends in Heterogeneous Data and a Proposal for New Assessment Methods. Anesth Analg. 2016;123(3):739-748. doi:10.1213/ANE.0000000000001389.

92 Jansson, C, Mittendorfer-Rutz, E & Alexanderson, K. Sickness absence because of musculoskeletal diagnoses and risk of all-cause and cause-specific mortality: a nationwide Swedish cohort study. Pain. 2012;153(5):998-1005. doi:10.1016/j.pain.2012.01.028.

93 Johnsen, LG et al. Comparison of the SF6D, the EQ5D, and the oswestry disability index in patients with chronic low back pain and degenerative disc disease. BMC Musculoskelet Disord. 2013;14:148. doi:10.1186/1471-2474-14-148.

94 Johnson, KB et al. Precision Medicine, AI, and the Future of Personalized Health Care. Clin Transl Sci. 2021;14(1):86-93. doi:10.1111/cts.12884.

95 Kaiser, U et al. Developing a core outcome domain set to assessing effectiveness of interdisciplinary multimodal pain therapy: the VAPAIN consensus statement on core outcome domains. Pain. 2018;159(4):673-683. doi:10.1097/j.pain.0000000000001129.

96 Kaiser, U, Treede, RD & Sabatowski, R. Multimodal pain therapy in chronic noncancer pain-gold standard or need for further clarification? Pain. 2017;158(10):1853-1859. doi:10.1097/j.pain.0000000000000902.

97 Kamper, SJ et al. Multidisciplinary biopsychosocial rehabilitation for chronic low back pain: Cochrane systematic review and meta-analysis. BMJ. 2015;350:h444. doi:10.1136/bmj.h444.

98 Karlsson, NE, Carstensen, JM, Gjesdal, S & Alexanderson, KA. Risk factors for disability pension in a population-based cohort of men and women on long-term sick leave in Sweden. Eur J Public Health. 2008;18(3):224-231. doi:10.1093/eurpub/ckm128.

99 Kelly, GA, Blake, C, Power, CK, O'Keeffe, D & Fullen, BM. The association between chronic low back pain and sleep: a systematic review. Clin J Pain. 2011;27(2):169-181. doi:10.1097/AJP.0b013e3181f3bdd5.

100 Kivimaki, M et al. Sickness absence as a risk marker of future disability pension: the 10-town study. J Epidemiol Community Health. 2004;58(8):710-711. doi:10.1136/jech.2003.015842.

101 Koller, D & Friedman, N. Probabilistic graphical models: principles and techniques. MIT press, 2009.

102 Kuijer, W, Groothoff, JW, Brouwer, S, Geertzen, JH & Dijkstra, PU. Prediction of sickness absence in patients with chronic low back pain: a systematic review. J Occup Rehabil. 2006;16(3):439-467. doi:10.1007/s10926-006-9021-8.

103 Laaksonen, M & Gould, R. Regional differences in disability retirement: explaining between-county differences in Finland. Scand J Work Environ Health. 2013;39(6):609-617. doi:10.5271/sjweh.3373.

104 Lash, TL, Fox, MP & Fink, AK. Applying Quantitative Bias Analysis to Epidemiologic Data. New York; Springer, 2009.

105 Latremoliere, A & Woolf, CJ. Central sensitization: a generator of pain hypersensitivity by central neural plasticity. J Pain. 2009;10(9):895-926. doi:10.1016/j.jpain.2009.06.012.

106 Leadley, RM, Armstrong, N, Lee, YC, Allen, A & Kleijnen, J. Chronic diseases in the European Union: the prevalence and health cost implications of chronic pain. J Pain Palliat Care Pharmacother. 2012;26(4):310-325. doi:10.3109/15360288.2012.736933.

107 Lewis, D. Counterfactuals. Oxford; Blackwell Publishers, 1973.

108 Lidwall, U. Termination of sickness benefits or transition to disability pension after changes in sickness insurance: a Swedish register study. Disabil Rehabil. 2013;35(2):118-124. doi:10.3109/09638288.2012.689920.

109 Lidwall, U. Sick leave diagnoses and return to work: a Swedish register study. Disabil Rehabil. 2015;37(5):396-410. doi:10.3109/09638288.2014.923521.

110 LoMartire, R, Constan, L & Frumento, P. Authors' reply to the comment by Rivano-Fischer and Stålnacke. Eur J Pain. 2021;In press. doi:10.1002/ejp.1890.

111 Ludvigsson, JF et al. External review and validation of the Swedish national inpatient register. BMC Public Health. 2011;11:450. doi:10.1186/1471-2458-11-450.

112 Ludvigsson, JF, Otterblad-Olausson, P, Pettersson, BU & Ekbom, A. The Swedish personal identity number: possibilities and pitfalls in healthcare and medical research. Eur J Epidemiol. 2009;24(11):659-667. doi:10.1007/s10654-009-9350-y.

113 Ludvigsson, JF, Svedberg, P, Olen, O, Bruze, G & Neovius, M. The longitudinal integrated database for health insurance and labour market studies (LISA) and its use in medical research. Eur J Epidemiol. 2019;34(4):423-437. doi:10.1007/s10654-019-00511-8.

114 Maetzel, A & Li, L. The economic burden of low back pain: a review of studies published between 1996 and 2001. Best Pract Res Clin Rheumatol. 2002;16(1):23-30. doi:10.1053/berh.2001.0204.

115 Mansournia, MA, Higgins, JP, Sterne, JA & Hernan, MA. Biases in Randomized Trials: A Conversation Between Trialists and Epidemiologists. Epidemiology. 2017;28(1):54-59. doi:10.1097/EDE.0000000000000564.

116 Mastekaasa, A & Melsom, A. Occupational segregation and gender differences in sickness absence: Evidence from 17 European countries. Eur Sociol Rev. 2014;30(5):582–594.

117 Maydeu-Olivares, A. Goodness-of-fit assessment of item response theory models. Measurement. 2013;11:71–101.

118 Meira-Machado, L, de Uña-Álvarez, J & Cadarso-Suárez, C. Multi-state models for the analysis of time-to-event data. Stat Methods Med Res. 2009;18(2):195-222.

119 Melzack, R. From the gate to the neuromatrix. Pain. 1999;Suppl 6:S121-S126. doi:10.1016/S0304-3959(99)00145-1.

120 Melzack, R. Evolution of the neuromatrix theory of pain. The Prithvi Raj Lecture: presented at the third World Congress of World Institute of Pain, Barcelona 2004. Pain Pract. 2005;5(2):85-94. doi:10.1111/j.1533-2500.2005.05203.x.

121 Merskey, H & Bogduk, N. Classification of Chronic Pain: Descriptions of Chronic Pain Syndromes and Definitions of Pain Terms. Seattle, WA, IASP Press, 1994.

122 Moll, LT et al. Return to Work in Employees on Sick Leave due to Neck or Shoulder Pain: A Randomized Clinical Trial Comparing Multidisciplinary and Brief Intervention with One-Year Register-Based Follow-Up. J Occup Rehabil. 2018;28(2):346-356. doi:10.1007/s10926-017-9727-9.

123 Moscelli, G, Siciliani, L, Gutacker, N & Cookson, R. Socioeconomic inequality of access to healthcare: Does choice explain the gradient? J Health Econ. 2018;57:290-314. doi:10.1016/j.jhealeco.2017.06.005.

124 National Board of Health and Welfare. The Swedish Prescribed Drug Register. url: https://www.socialstyrelsen.se/en/statistics-and-data/registers/register-information/the-swedish-prescribed-drug-register/. Accessed: 2020.

125 National Board of Health and Welfare. The National Patient Register. url: https://www.socialstyrelsen.se/en/statistics-and-data/registers/register-information/the-national-patient-register/. Accessed: 2020.

126 National Board of Health and Welfare. [Uppföljning av primärvård och omställningen till en mer nära vård. Deluppdrag I – Nationell insamling av registeruppgifter från primärvården]. 2021. Report No: 2021-2-7223.

127 Nicholas, M et al. The IASP classification of chronic pain for ICD-11: chronic primary pain. Pain. 2019;160(1):28-37. doi:10.1097/j.pain.0000000000001390.

128 Niv, D & Devor, M. Chronic Pain as a Disease in its Own Right. Pain Practice. 2004;4(3):179-181.

129 Norlund, A, Ropponen, A & Alexanderson, K. Multidisciplinary interventions: review of studies of return to work after rehabilitation for low back pain. J Rehabil Med. 2009;41(3):115-121. doi:10.2340/16501977-0297.

130 Norton, S, Cosco, T, Doyle, F, Done, J & Sacker, A. The Hospital Anxiety and Depression Scale: a meta confirmatory factor analysis. J Psychosom Res. 2013;74(1):74-81. doi:10.1016/j.jpsychores.2012.10.010.

131 Nyberg, V, Sanne, H & Sjolund, BH. Swedish quality registry for pain rehabilitation: purpose, design, implementation and characteristics of referred patients. J Rehabil Med. 2011;43(1):50-57. doi:10.2340/16501977-0631.

132 Osterlund, N. [MiDAS: Sjukpenning och Rehabiliteringspenning Verson 1.02]. Stockholm, Swedish Social Insurance Agency, 2011.

133 Pallant, JF & Bailey, CM. Assessment of the structure of the Hospital Anxiety and Depression Scale in musculoskeletal patients. Health Qual Life Outcomes. 2005;3:82. doi:10.1186/1477-7525-3-82.

134 Pallant, JF & Tennant, A. An introduction to the Rasch measurement model: an example using the Hospital Anxiety and Depression Scale (HADS). Br J Clin Psychol. 2007;46(Pt 1):1-18. doi:10.1348/014466506x96931.

135 Patel, AS et al. The impact and burden of chronic pain in the workplace: a qualitative systematic review. Pain Pract. 2012;12(7):578-589. doi:10.1111/j.1533-2500.2012.00547.x.

136 Pearl, J. Causality: models, reasoning, and inference. New York; Cambridge University Press, 2009.

137 Pedersen, P, Nielsen, CV, Jensen, OK, Jensen, C & Labriola, M. Employment status five years after a randomised controlled trial comparing multidisciplinary and brief intervention in employees on sick leave due to low back pain. Scand J Public Health. 2018;46(3):383-388. doi:10.1177/1403494817722290.

138 Pedregosa, F et al. Scikit-learn: Machine learning in python. J Mach Learn Res. 2011;12:2825-2830.

139 Phillips, CJ. The Cost and Burden of Chronic Pain. Rev Pain. 2009;3(1):2-5. doi:10.1177/204946370900300102.

140 Piccarreta, R & Studer, M. Holistic analysis of the life course: Methodological challenges and new perspectives. Adv Life Course Res. 2019;41.

141 Prefontaine, K & Rochette, A. A Literature Review on Chronic Pain: The Daily Overcoming of a Complex Problem. 2013;76(6):280-286. doi:10.4276/030802213x13706169932905.

142 Prieto, L et al. A Rasch model analysis to test the cross-cultural validity of the EuroQoL-5D in the Schizophrenia Outpatient Health Outcomes Study. Acta Psychiatr Scand Suppl. 2003(416):24-29. doi:10.1034/j.1600-0447.107.s416.6.x.

143 Prieto, L & Sacristán, JA. What is the value of social values? The uselessness of assessing health-related quality of life through preference measures. BMC Med Res Methodol. 2004;4(1):1-9.

144 Python Software Foundation. Python Language Reference (version 3.8.3). 2020. url: http://www.python.org.

145 R Core Team. R: A language and environment for statistical computing (versions 3.5.2 and 4.0.2). R Foundation for Statistical Computing, 2020. url: https://www.R-project.org/.

146 Raffaeli, W & Arnaudo, E. Pain as a disease: an overview. J Pain Res. 2017;10:2003-2008. doi:10.2147/JPR.S138864.

147 Raja, SN et al. The revised International Association for the Study of Pain definition of pain: concepts, challenges, and compromises. Pain. 2020. doi:10.1097/j.pain.0000000000001939.

148 Rayner, L et al. Depression in patients with chronic pain attending a specialised pain treatment centre: prevalence and impact on health care costs. Pain. 2016;157(7):1472-1479. doi:10.1097/j.pain.0000000000000542.

149 Reckase, M. Multidimensional item response theory. New York; Springer, 2009.

150 Region Skane. [Smärtrehabilitering: Nationellt kvalitetsregister över Smärtrehabilitering (NRS)]. url: https://skr.se/kvalitetsregister/hittaregister/registerarkiv/smartrehabilitering. Accessed: 2019.

151 Reise, S & Revicki, D. Handbook of item response theory modeling: applications to typical performance assessment. New York; Routledge, 2015.

152 Ritschard, G & Studer, M. Sequence Analysis and Related Approaches: Innovative Methods and Applications. Cham, Switzerland; Springer, 2018.

153 Rivano-Fischer, MR et al. Variability in patient characteristics and service provision of interdisciplinary pain rehabilitation: A study using the Swedish national quality registry for pain rehabilitation. J Rehabil Med. 2020;52(11):jrm00128. doi:10.2340/16501977-2765.

154 Rivano-Fischer, MR & Stalnacke, B. Comment on `The value of interdisciplinary treatment for sickness absence in chronic pain: A nationwide register-based cohort study`. Eur J Pain. 2021;In press. doi:10.1002/ejp.1873.

155 Roche-Leboucher, G et al. Multidisciplinary intensive functional restoration versus outpatient active physiotherapy in chronic low back pain: a randomized controlled trial. Spine (Phila Pa 1976). 2011;36(26):2235-2242. doi:10.1097/BRS.0b013e3182191e13.

156 Ropponen, A et al. Predicting the duration of sickness absence spells due to back pain: a population-based study from Sweden. Occup Environ Med. 2020;77(2):115-121. doi:10.1136/oemed-2019-106129.

157 Rothman, KJ, Greenland, S & Lash, TL. Modern epidemiology. Philadelphia; Wolters Kluwer Health/Lippincott Williams & Wilkins, 2008.

158 Rubin, DB. Estimating causal effects of treatments in randomized and nonrandomized studies. J Educ Psychol. 1974;66:688–701.

159 Samejima, F. Estimation of latent ability using a response pattern of graded scores. Psychometrika Monogr. 1969;34:1–97.

160 Scascighini, L, Toma, V, Dober-Spielmann, S & Sprott, H. Multidisciplinary treatment for chronic pain: a systematic review of interventions and outcomes. Rheumatology (Oxford). 2008;47(5):670-678. doi:10.1093/rheumatology/ken021.

161 Shadish, WR, Cook, TD & Campbell, DT. Experimental and Quasi-Experimental Designs for Generalized Causal Inference. Houghton, Mifflin and Company. Boston; Houghton, Mifflin and Company, 2002.

162 Shmueli, G. To Explain or to Predict? Stat Sci. 2010;25:289-310.

163 Statistics Sweden. https://www.scb.se/LISA. Accessed: 2019.

164 Steers, RM & Rhodes, SR. Major influences on employee attendance: a process model. J Appl Psychol. 1978;63(4):391-407.

165 Steingrimsdottir, OA, Landmark, T, Macfarlane, GJ & Nielsen, CS. Defining chronic pain in epidemiological studies: a systematic review and meta-analysis. Pain. 2017;158(11):2092-2107. doi:10.1097/j.pain.0000000000001009.

166 Stochl, J et al. Usefulness of EQ-5D for evaluation of health-related quality of life in young adults with first-episode psychosis. Qual Life Res. 2013;22(5):1055-1063. doi:10.1007/s11136-012-0222-7.

167 Stover, M et al. Unemployment and disability pension--an 18-year follow-up study of a 40-year-old population in a Norwegian county. BMC Public Health. 2012;12:148. doi:10.1186/1471-2458-12-148.

168 Strom, P et al. Artificial intelligence for diagnosis and grading of prostate cancer in biopsies: a population-based, diagnostic study. Lancet Oncol. 2020;21(2):222-232. doi:10.1016/S1470-2045(19)30738-7.

169 Studer, M. WeightedCluster Library Manual: A practical guide to creating typologies of trajectories in the social sciences with R. LIVES Working Papers. 2013:1-32.

170 Sveriges Kommuner och Regioner. Nationella kvalitetsregister. url: https://skr.se/kvalitetsregister/. Accessed: Sep 2021.

171 Swedish Agency for Health Technology Assessment and Assessment of Social Services. [Metoder för behandling av långvarig smärta. En systematisk litteraturöversikt]. Stockholm, 2006. Report No: 91-85413-08-9.

172 Swedish Agency for Health Technology Assessment and Assessment of Social Services. [Rehabilitering vid långvarig smärta. En systematisk litteraturöversikt]. Stockholm, 2010. Report No: 198.

173 Swedish Social Insurance Agency. [Vad kostar olika sjukdomar i sjukförsäkringen? Kostader för sjukpenning i sjukskrivningar (över 14 dagar) samt sjukersättning och aktivitetsersättning år 2009 fördelat på diagnos: Socialförsäkringsrapport 2011:4]. 2011.

174 Swedish Social Insurance Agency. [Sjukfrånvarons utveckling 2016 (Socialförsäkringsrapport 2016:7)]. Stockholm, Sweden, Swedish Social Insurance Agency, 2016.

175 Swedish Social Insurance Agency. [Ohälsotalet efter kommun och ålder, 2001−2019]. url: https://www.forsakringskassan.se/statistik/sjuk/ohalsomatt. Accessed: 2020.

176 Syriopoulou, E, Mozumder, SI, Rutherford, MJ & Lambert, PC. Robustness of individual and marginal model-based estimates: A sensitivity analysis of flexible parametric models. Cancer Epidemiol. 2019;58:17-24. doi:10.1016/j.canep.2018.10.017.

177 Taft, C, Karlsson, J & Sullivan, M. Do SF-36 summary component scores accurately summarize subscale scores? Qual Life Res. 2001;10(5):395-404. doi:10.1023/a:1012552211996.

178 Task Force on Taxonomy of the International Association for the Study of Pain. Classification of chronic pain: descriptions of chronic pain syndromes and definitions of pain terms. Seattle; IASP Press, 1994.

179 Taylor, P. Sickness Absence: Facts and Misconceptions. J R Coll Physicians Lond. 1974;8(4):315-333.

180 Terwee, CB et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007;60(1):34-42. doi:10.1016/j.jclinepi.2006.03.012.

181 Titman, AC & Sharples, LD. Model diagnostics for multi-state models. Stat Methods Med Res. 2010;19:621–651.

182 Toye, F, Seers, K, Hannink, E & Barker, K. A mega-ethnography of eleven qualitative evidence syntheses exploring the experience of living with chronic non-malignant pain. BMC Med Res Methodol. 2017;17(1):116. doi:10.1186/s12874-017-0392-7.

183 Treede, RD et al. Chronic pain as a symptom or a disease: the IASP Classification of Chronic Pain for the International Classification of Diseases (ICD-11). Pain. 2019;160(1):19-27. doi:10.1097/j.pain.0000000000001384.

184 Tucker, G, Adams, R & Wilson, D. Observed agreement problems between sub-scales and summary components of the SF-36 version 2 - an alternative scoring method can correct the problem. PLoS One. 2013;8(4):e61191. doi:10.1371/journal.pone.0061191.

185 Turk, DC et al. Core outcome domains for chronic pain clinical trials: IMMPACT recommendations. Pain. 2003;106(3):337-345. doi:10.1016/j.pain.2003.08.001.

186 van Middelkoop, M et al. A systematic review on the effectiveness of physical and rehabilitation interventions for chronic non-specific low back pain. Eur Spine J. 2011;20(1):19-39. doi:10.1007/s00586-010-1518-3.

187 van Tulder, MW, Koes, BW & Bouter, LM. A cost-of-illness study of back pain in The Netherlands. Pain. 1995;62(2):233-240. doi:10.1016/0304-3959(94)00272-G.

188 Virtanen, M et al. Lifestyle factors and risk of sickness absence from work: a multicohort study. Lancet Public Health. 2018;3(11):e545-e554. doi:10.1016/S2468-2667(18)30201-9.

189 Virtanen, P et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat Methods. 2020;17(3):261-272. doi:10.1038/s41592-019-0686-2.

190 Wallerstedt, SM, Wettermark, B & Hoffmann, M. The First Decade with the Swedish Prescribed Drug Register - A Systematic Review of the Output in the Scientific

Literature. Basic Clin Pharmacol Toxicol. 2016;119(5):464-469. doi:10.1111/bcpt.12613.

191 Wallman, T et al. Sick-leave track record and other potential predictors of a disability pension. A population based study of 8,218 men and women followed for 16 years. BMC Public Health. 2009;9:104. doi:10.1186/1471-2458-9-104.

192 Ware, JE, Jr. et al. The factor structure of the SF-36 Health Survey in 10 countries: results from the IQOLA Project. International Quality of Life Assessment. J Clin Epidemiol. 1998;51(11):1159-1165. doi:10.1016/s0895-4356(98)00107-3.

193 Ware, JE, Snow, KK, Kosinski, M & Gandek, B. SF-36 health survey: manual and interpretation guide. Lincoln; Quality Metric, 1993.

194 Wettermark, B et al. The new Swedish Prescribed Drug Register--opportunities for pharmacoepidemiological research and experience from the first six months. Pharmacoepidemiol Drug Saf. 2007;16(7):726-735. doi:10.1002/pds.1294.

195 Willekens, F. Multistate analysis of life histories with R. Springer, 2014.

196 Williamson, T. The Philosophy of Philosophy. Oxford; Blackwell Publishing, 2007.

197 World Health Organization. International classification of diseases for mortality and morbidity statistics (11th Revision). url: https://icd.who.int. Accessed: 2018.

198 Zigmond, AS & Snaith, RP. The hospital anxiety and depression scale. Acta Psychiatr Scand. 1983;67(6):361-370. doi:10.1111/j.1600-0447.1983.tb09716.x.

10 APPENDIX

10.1 STATISTICAL SOFTWARE

Table 1S. R packages

Study 1		Study 2		Study 3	
Name	**Version**	**Name**	**Version**	**Name**	**Version**
car	3.0-3	corrplot	0.84	data.table	1.13.0
corrplot	0.84	cowplot	1.0.0	DescTools	0.99.37
ggplot2	3.2.0	data.table	1.13.0	flexsurv	1.1.1
ggExtra	0.8	DescTools	0.99.37	ggplot2	3.3.2
ggpubr	0.2.1	dplyr	1.0.2	haven	2.3.1
gridExtra	2.2.1	ggalluvial	0.12.2	ipw	1.0-11
lemon	0.4.3	ggplot2	3.3.2	lubridate	1.7.9
lubridate	1.7.4	haven	2.3.1	msSurv	1.2-2
mirt	1.30	Hmisc	4.4-1	mstate	0.2.12
plyr	1.8.4	IRanges	2.22.2	rstpm2	1.5.1
scales	1.0.0	lubridate	1.7.9	scales	1.1.1
		psych	2.0.7	stringr	1.4.0
		stringr	1.4.0	survival	3.1-12
		tidyverse	1.3.0	tidyverse	1.3.0
		TraMineR	2.2-0.1	TraMineR	2.2-0.1
		WeightedCluster	1.4-1		

R version 3.5.2 was used in study 1, while R version 4.0.2 was used in studies 2 and 3.

Table 2S. Python libraries used in study 2.

Name	Version
eli5	0.10.1
IPython	7.16.1
matplotlib	3.2.2
mlxtend	0.18.0
numpy	1.18.5
pandas	1.0.5
pdpbox	0.2.0
pickle	0.7.5
pycm	2.5
scipy	1.5.0
seaborn	0.10.1
sklearn	0.23.2
tensorflow	2.4.0
xgboost	1.2.0

Python version 3.8.3.